

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2002, 17:42:59 ; Search time 34.93 Seconds

(without alignments)
2314.965 Million cell updates/sec

Title: us-09-423-516-2

Perfect score: 4126
Sequence: 1 MWYTKLIPALLQHVLHLL.....VAYYAKWIKHILTYKVPQS 728

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Cal number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_032802:*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4126	100.0	728	14 AAR42062	Vascular endothelial
2	4126	100.0	728	17 AAM00338	Human hepatic paren
3	4126	100.0	728	17 AAM00340	Wild type hepatocyte
4	4126	100.0	728	19 AAW59922	Human leukocyte-de
5	4126	100.0	728	19 AAW58966	Human hepatocyte g
6	4126	100.0	728	19 AAW42998	Recombinant human
7	4126	100.0	728	19 AAM39207	Human hepatocyte g
8	4126	100.0	728	14 AAR39521	Hepatocyte growth
9	4119	99.8	728	13 AAR25676	Recombinant human
10	4119	99.8	728	14 AAR40862	Competative inhibi
11	4118	99.8	728	13 AAR20005	Human hepatocyte g

12	4116	99.8	728	20 AAM88529	Human hepatocyte g
13	4115	99.7	727	12 AAR10656	Hepatic parenchyma
14	4113	99.7	728	20 AAM88532	Human hepatocyte g
15	4112	99.7	728	12 AAR15623	Human leukocyte-de
16	4112	99.7	728	20 AAM88531	Human hepatocyte g
17	4111	99.6	728	20 AAM88530	Human hepatocyte g
18	4110	99.6	728	17 AAR87525	Human hepatocyte
19	4109	99.6	728	17 AAR87523	Mutant hepatocyte
20	4108	99.6	728	14 AAR40863	Mutant hepatocyte
21	4107	99.5	728	14 AAR14243	Competative inhibi
22	4106	99.5	728	17 AAR87524	Human hepatocyte g
23	4096	99.3	728	15 AAR47227	Mutant hepatocyte
24	4095	99.2	728	15 AAR87522	Hepatocyte growth
25	4088.5	99.1	723	11 AAR07144	Tumour cytotoxic f
26	4088.5	99.1	723	12 AAR15624	Human leukocyte-de
27	4088.5	99.1	723	15 AAR57026	Human wild-type tu
28	4088.5	99.1	723	16 AAR82685	Human cytotoxic f
29	4088.5	99.1	723	16 AAR76690	Tumour cytotoxic f
30	4085.5	99.0	723	13 AAR21142	Human plasmalogen-
31	4085.5	99.0	723	15 AAR57027	Human TCF-II. Hom
32	4085.5	99.0	723	15 AAR57028	Human modified tum
33	4081.5	98.9	723	13 AAR25677	Human modified tum
34	4078.5	98.8	723	19 AAR59923	Recombinant human
35	4073	98.7	728	13 AAR25160	Human leukocyte-de
36	4066	98.5	728	13 AAR12792	Human HGF. Homo s
37	4066	98.5	728	19 AAW48221	Human hepatocyte g
38	4066	98.5	728	20 AAW97041	Human recombinant
39	4066	98.5	728	22 AAG68090	A human hepatocyte
40	4066	98.5	728	22 AAB84517	Human hepatocyte g
41	4064.5	98.5	723	13 AAR29819	Amilo acid sequenc
42	4064.5	98.5	723	17 AAR99688	TCF-II. AAR29819
43	4064.5	98.5	723	17 AAR99689	TCF mutant having
44	4061	98.4	728	13 AAR21976	Human hepatocyte g
45	4061	98.4	728	14 AAR52942	Human hepatocyte g

ALIGNMENTS

RESULT	1	
AAR42062	1	
ID	AAR42062	standard; Protein; 728 AA.
XX		
AC	AAR42062;	
XX		
DT	15-NOV-1993	(first entry)
XX		
DE	Vascular endothelial cell growth enhancer protein.	
XX		
KW	Enhance; growth; vascular endothelial cell; human; tumour; cell line;	
KW	HUCCA-II; HUCCA-III; blood vessel; wounds; burns; decubitus;	
KW	post-operative tissue damage; drug; cardiac angiopathy.	
XX		
OS	Homo sapiens.	
XX		
PN	EP550296-A.	
XX		
PD	07-JUL-1993.	
XX		
PF	27-NOV-1992;	92EP-0403199.
XX		
PR	28-NOV-1991;	91JP-0337999.
XX		
PA	(TERU) TERUMO CORP.	
XX		
PI	Adachi M, Harada K, Hirahara I, Sudo T;	
XX		
DR	WPI; 1993-215669/27.	
XX		
DR	N-PSDB; AAQ45702.	
XX		
PT	Vascular endothelial cell growth factor protein - used for	
PT	promoting angiogenesis in the treatment of cardiac angiopathy,	
PT	wounds, burn injuries, postoperative tissue damage etc.	

XX Claim 6; Page 18-21; 44pp; English.
XX
XX This sequence represents a single chain protein which selectively
CC enhances the growth of vascular endothelial cells. This protein
CC was produced by the human tumour cell line HUCCA-II or HUCCA-III.
CC This protein enhances the formation of new blood vessels and may be
CC used to enhance healing of wounds, burns injuries, decubitus or
CC post-operative tissue damage. It may also be used as a drug for
CC cardiac angiopathy.
XX
XX
SQ Sequence 728 AA;

Query Match 100.0%; Score 4126; DB 14; Length 728;
Best Local Similarity 100.0%; Pred. No. 9.5e-277;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTVTLPLALLQHVLLHLLELPALPYAEGQRRKRRNTIHEFKSAKTTLIKIDPALKIK 60
1 mmtvtkllpallllqhvllhlhlllpiaipyaeqgrrkrrntihfkksaktllikidpalkik 60
61 TKKVNADQCANRCTRNKGLPFTCKAFVBDKARKQCLMPFPNSMSGYKKEFGHEFDLYE 120
61 tkkvnadqcanrcrnkglpftckafvbdkarkqclwfpnsmssgykkefghfdlye 120
DB 121 NKDYIRNCIIIGKRSYKGYSTIRKSGIKQOPWSSMIPEHSFLPSSYRGKDLQENYCRNP 180
121 nkdyirnciigkrsykygystirksgikcqpwsmlpeshflpssyrgkdlqenyccrnp 180
QY 181 RGEEGPWCFTSNPEVRYEVCDDIPQCEVECMTCNGESYRGMLDHTESGKICQRDHOQTP 240
181 rgeegpwcftsnpevryevcdipqcevecmtcngesyrgmlhtesgkicqrwdhoqtp 240
DB 181 RGEEGPWCFTSNPEVRYEVCDDIPQCEVECMTCNGESYRGMLDHTESGKICQRDHOQTP 240
181 rgeegpwcftsnpevryevcdipqcevecmtcngesyrgmlhtesgkicqrwdhoqtp 240
QY 241 HRHAKLPERYEDKGGDDNYCRNPDCQPRPWCYTLDPHRMWECALIKTCADNMNTDVAL 300
241 hrhakeperyedkggddnycrnpdcqprpwcytldphrmwecaliktcadnmntdval 300
DB 241 hrhakeperyedkggddnycrnpdcqprpwcytldphrmwecaliktcadnmntdval 300
QY 301 ETTETICQGGEGYRGTVNTIMNGICQWRDSQYRPHEDMTPEFKCKDLRENYCRNPDS 360
301 ettetiqgggyrgtvntimngicqwrdsqyrphedmtpefkckdlrenyccrnpds 360
DB 301 ettetiqgggyrgtvntimngicqwrdsqyrphedmtpefkckdlrenyccrnpds 360
QY 361 ESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGILSOTRSGLTCSMDKME 420
361 espwcttdpnirvgycsqipncdmshgdcyrgngknymgilsotrsgltcsmdkme 420
DB 361 espwcttdpnirvgycsqipncdmshgdcyrgngknymgilsotrsgltcsmdkme 420
QY 421 DLHRIIFMRPDSKINENYCRNPDDAHGPMCYTGMPCLIPMPYCPISRCESGTPPIIVL 480
421 dlhriifmrpdsकिननय्क्रनपद्दहगपम्यत्गमप्लिपम्यचपिसरसेगत्पपििवल 480
DB 421 dlhriifmrpdsकिननय्क्रनपद्दहगपम्यत्गमप्लिपम्यचपिसरसेगत्पपििवल 480
QY 481 DHPVISCATKQLRVNGIPIRTNIGMWVSLRYRNKHICGSLIKESVWLTAROCPPSD 540
481 dhpviscatkqlrvngipirtnigmwvslryrnkhicgslikesvwltaarocppsd 540
DB 481 dhpviscatkqlrvngipirtnigmwvslryrnkhicgslikesvwltaarocppsd 540
QY 541 LKDYPAWLGIDHVGGRDECKQVNLVNSQVYGPBGSDLVLMKLARPAVLDFVSTIDLP 600
541 lkdy pawlgidhvggrdeckqvnlvnsqvypbgsvdlvmlklarpavldfvstidlp 600
DB 541 lkdy pawlgidhvggrdeckqvnlvnsqvypbgsvdlvmlklarpavldfvstidlp 600
QY 601 NYGCTIIPKTCSSVYGMGTGLINVDGLIRVAHLVIMGNEKSOHQHGRGVTINSEITCG 660
601 nygctiipktcssvygmgtglinvdglirvahlvimgneksohqhgrgvtinseitcg 660
DB 601 nygctiipktcssvygmgtglinvdglirvahlvimgneksohqhgrgvtinseitcg 660
QY 661 AEKISSGCEGDYGPPLVCEQHKRMWLVGVIVPGRGCAIPNPGIFVRAVAYAKWHIKII 720
661 aekissgcegdypplvceqhkrmwlvgvivpgrgcaipnpgifvrayayakwhikii 720
DB 661 aekissgcegdypplvceqhkrmwlvgvivpgrgcaipnpgifvrayayakwhikii 720
QY 721 LTYKVPQS 728
721 ltykvps 728
DB 721 ltykvps 728

RESULT 2
AAM00338

ID AAM00338 standard; protein; 728 AA.
XX
XX AAM00338;
AC
XX 27-JUN-1997 (first entry)
XX
XX Human hepatic parenchymal cell growth factor.
DE
XX Human hepatic parenchymal cell; growth factor; HGF; treatment;
KW prevention; ischemia; ischemic; disease; reperfusion; disorder;
KW blood; liver; transplant; acute; failure; ischemia; ischemic;
KW organ.
XX
XX Homo sapiens.
OS
XX W09632960-A1.
PN
XX 24-OCT-1996.
PD
XX 19-APR-1996; 96WO-JP01065.
PE
XX 21-APR-1995; 95JP-0096994.
PR
XX (MITU) MITSUBISHI CHEM CORP.
PA
XX Gemba M, Yonehana T;
PI
XX WPI; 1996-48555/48.
DR
XX Hepatic parenchymal cell growth factors - used as remedy and/or
PT preventive medicine for ischemic diseases
PS
XX Claim 9; Pages 7-10; 16pp; Japanese.
XX
XX The present sequence is the human hepatic parenchymal cell growth
CC factor (HGF), which has a molecular weight of 76-92 kD estimated by
CC SDS-PAGE, parenchymal cell growth activity, which is lost following
CC heat treatment at 80 degrees C for 10 minutes and strong affinity
CC towards heparin. It may be used to treat and prevent ischemic
CC diseases, e.g. blood reperfusion disorder, ischemic reperfusion
CC disorders in liver transplant, acute liver failure and organ
CC disorder during organ transplant. For adults, it is administered at
CC a level of 1 microg to 10 mg/kg/day, preferably 10-1000
CC microg/kg/day.
CC A pig liver derived cultured epithelial cell line was cultured in
CC DMEM, cow embryo serum and N-2-hydroxyethylpiperazine ethane
CC sulphonic acid to confluence. The medium was then changed to serum
CC and glucose free DMEM, and the oxygen concentration reduced to less
CC than 2%. The cells were cultured for 6 hours, and then for another
CC hour at 95% air/5% carbon dioxide (reoxygenation). HGF was added
CC just before hypoxiation or reoxygenation. HGF suppressed free
CC lactic acid dehydrogenase (LDH) by 49.6% when added before
CC hypoxiation, and by 60.9% when added before reoxygenation, at 50
CC ng/ml (free LDL was 100% in the absence of HGF).
XX
SQ Sequence 728 AA;

Query Match 100.0%; Score 4126; DB 17; Length 728;
Best Local Similarity 100.0%; Pred. No. 9.5e-277;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTVTLPLALLQHVLLHLLELPALPYAEGQRRKRRNTIHEFKSAKTTLIKIDPALKIK 60
1 mmtvtkllpallllqhvllhlhlllpiaipyaeqgrrkrrntihfkksaktllikidpalkik 60
DB 1 mmtvtkllpallllqhvllhlhlllpiaipyaeqgrrkrrntihfkksaktllikidpalkik 60
QY 61 TKKVNADQCANRCTRNKGLPFTCKAFVBDKARKQCLMPFPNSMSGYKKEFGHEFDLYE 120
61 tkkvnadqcanrcrnkglpftckafvbdkarkqclwfpnsmssgykkefghfdlye 120
DB 61 tkkvnadqcanrcrnkglpftckafvbdkarkqclwfpnsmssgykkefghfdlye 120
QY 121 NKDYIRNCIIIGKRSYKGYSTIRKSGIKQOPWSSMIPEHSFLPSSYRGKDLQENYCRNP 180
121 nkdyirnciigkrsykygystirksgikcqpwsmlpeshflpssyrgkdlqenyccrnp 180
DB 121 nkdyirnciigkrsykygystirksgikcqpwsmlpeshflpssyrgkdlqenyccrnp 180

```

QY 181 RGEGRPWCFSTNPEVRYEVDIPQCESEVEMTCNGESYRGIMDHTESGKICQRMHQTP 240
    |||||||
Db 181 rgeegpwtctsnpevryevcdipqcevecmctcngesyrgimhtesgkicqwmhqtpr 240
QY 241 HRRKFLPERYPDKGFDNDYCRNPDGPRPWCYTLDPHTRMETCAIKTCADMTMDTVPL 300
    |||||||
Db 241 hrnkflperypdkgfdndyrcnpgdprpwcyltldphtremetcaiktcadmtmdtvp1 300
QY 301 ETRRCTOGGEGYRGNTNTWNGIPCORMDSQYRPHEDMTPEMKCKDLNENYCRNDGS 360
    |||||||
Db 301 etrec1ggggyrgntnltnwngipcgwdsqyphedmtpepkckdlenyrcnpgds 360
QY 361 ESPWCFETDNRIRYGCYQIPNCDMSHGDCYRNGKNYMGNTLSQTSGLTCSWMDKME 420
    |||||||
Db 361 espwcfetdnirvygcsqipncdmsngdcyrgnknymgnlsqtsgltsmwdkme 420
QY 421 DLRRHIFWEPDASKLNENYCRNPDGPRPWCYTLDPHTRMETCAIKTCADMTMDTVPL 480
    |||||||
Db 421 dlrrhifwepdasklneycrnpddahpwcylgnp1lpwdyp1srcegdttplvnl 480
QY 481 DHPYISCAKTKQLRVNNGIFPTNIGMWSLRANKHICGSLIKESWVLTARCCFERSD 540
    |||||||
Db 481 dhpviscaktkqlrvnngifptnigwvs1rynkhi1cgss1ikeswvltarccfprsd 540
QY 541 LKDYEAWLGIHDYHGRGDECKOVLANSOLVYGEPSDLVLMKLARPAVLDDFVSTIDLP 600
    |||||||
Db 541 lkdyeawlg1hdvhygrgdekckqylnvsq1lypgegsdlv1mklarparvlddfvstldlp 600
QY 601 NYGCTIPEKTSYVYGMGYTGLINYDGLLRVLAHLIYMGNEKCSQHRGKVTYLNSEICAG 660
    |||||||
Db 601 nygctipektsesyvgytg1inydg1lrvahlyimgnekcsqhrhgkvtlinese1cag 660
QY 661 AERTIGSPGCGDVGGLVCSQHKRMVLYIVPGRGCAINRRPPIFPRVAYAAWIKII 720
    |||||||
Db 661 aertigspcgdygg9p1vpcsqhkmrmv1gvlvpg9ca1prrp1fvravayakw1khi1 720
QY 721 LTYKVPQS 728
    |||||||
Db 721 ltykvpqs 728

RESULT 3
ID AAM00340 standard; protein; 728 AA.
XX AAM00340;
XX AAM00340;
XX 09-DEC-1996 (first entry)
XX wild type hepatocyte growth factor.
XX Human: hepatocyte growth factor; HGF; huHGF; serum; proteolytic cleavage;
KW pro-hormone; beta subunit; alpha subunit; kring1e domain; prothrombin;
KM plasminogen; catalytic domain; serine protease; HGF variant;
KW HGF receptor; malignancy; chronic HGF receptor activation.
XX Homo sapiens.
XX OS
XX FH Key Location/Qualifiers
XX FT Cleavage-site 494..495
XX FT /label= "Proteolytic-cleavage-site"
XX FT /note= "generates alpha and beta subunits"
XX FT Disulfide-bond 487..604
XX FT Peptide 1..54
XX FT /note= "prepro-sequence"
XX FT Peptide 1..31
XX FT /note= "Hydrophobic signal peptide"
XX FT Domain 128..206
XX FT /label= kring1e_1_domain
XX FT Domain 211..288
XX FT /label= kring1e_2_domain
XX FT Domain 303..383

```

```

FT FT /label= kring1e_3_domain
FT FT 391..464
FT FT /label= kring1e_4_domain
FT FT 294
FT FT Modified-site
FT FT /label= N-1-linked-glycosylation_site
FT FT 402
FT FT Modified-site
FT FT /label= N-1-linked-glycosylation_site
FT FT 566
FT FT Modified-site
FT FT /label= N-1-linked-glycosylation_site
FT FT 653
FT FT /label= N-1-linked-glycosylation_site
FT FT 494
FT FT /note= "Position of opt. substitution, esp. Glu,
FT FT Asp or Asn"
FT FT Misc-difference 495
FT FT /note= "Position of opt. substitution, esp. Tyr or Phe"
FT FT Misc-difference 534
FT FT /note= "Position of opt. substitution"
FT FT Misc-difference 673
FT FT /note= "Position of opt. substitution"
FT FT Misc-difference 692
FT FT /note= "Position of opt. substitution"
XX XX US547856-A.
XX XX 20-AUG-1996.
XX XX 18-MAY-1992; 92US-0884811.
XX XX 13-JUL-1993; 93US-0087783.
XX XX 18-MAY-1992; 92US-0884811.
XX XX 18-MAY-1992; 92US-0885971.
XX XX (GETH ) GENENTECH INC.
XX XX Godowski PJ, Lokker NA, Mark MR;
XX XX WPI; 1996-392634/39.
XX XX New hepatocyte growth factor variants - are resistant to in vivo
XX XX proteolytic cleavage into a 2-chain form, useful as HGF antagonists
XX XX PS Disclosure; Column 51-56; 39pp; English.
XX XX This sequence represents the wild type full length sequence of human
XX XX hepatocyte growth factor (HGF). HGF is isolated from human serum
XX XX and is a disulphide linked heterodimer derived by proteolytic cleavage
XX XX of this pro-hormone form between residues 494 and 495. This generates a
XX XX molecule composed of an alpha-subunit of 440 amino acids (mol. wt. 69
XX XX kD) and a beta-subunit of 234 amino acids (mol. wt. 34 kD). The alpha
XX XX and beta subunits are encoded by a single open reading frame. The alpha
XX XX subunit contains four kring1e domains based on their homology to
XX XX kring1e-like domains in other proteins, e.g. prothrombin, plasminogen.
XX XX The beta subunit shows high homology to the catalytic domain of serine
XX XX proteases. However two of the three residues which form the catalytic
XX XX triad of serine proteases are not conserved in HGF. Therefore, the
XX XX precise function of the beta chain remains unknown. The invention
XX XX includes HGF variants which retain HGF receptor binding activity without
XX XX having the biological activity of wild-type HGF. They can be used for
XX XX the treatment of pathological conditions associated with the activation
XX XX of a HGF receptor such as malignancies associated with chronic HGF
XX XX receptor activation.
XX XX Sequence 728 AA;
XX XX SQ
QY 1 MWVTKLPALLQVHLHLPLPAIPAEQGRKRRNTHEFFKSATKTLIKIPAIKIK 60
Db 1 mvvtkl1pall1qvl1h1l1l1p1a1p1a1p1a1p1a1p1a1p1a1p1a1p1a1p1a1k 60
Query Match 100.0%; Score 4126; DB 17; Length 728;
Best Local Similarity 100.0%; Pred. No. 9, 5e-277;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	61	TVVNVADQCANNCTRNKKLPLEPTCAFAFVDPKAROKCLMFPPNSSSGKKKEFHEEDLYE	120
Db	61	UKVNVADQCANNCTRNKKLPLEPTCAFAFVDPKAROKCLMFPPNSSSGKKKEFHEEDLYE	120
QY	121	NKDYNINCILIGGRSYKGVSTTKGIGICOPMSSWIPHEHFLPSSRGMDLENCRN	180
Db	121	nkdyninciligrsykvgstltskglcqpssmshphsflpssrygkdleqncrn	180
QY	181	RGEEGPWCFTSNPEYREVECDIPCCSEVECMTCNGESYRGLMDHTESGKICQRMHDQTP	240
Db	181	rgeegpwcftsnpeyrevecdipccsevecmctngesyrglmdhtesgkicqrwdhqp	240
QY	241	HRHKFLPEEYPPKGFEDNCRMPDQRPKMYCTLDPHTRMYEYCAIKRCANTMNDVPL	300
Db	241	hrhkflpeyppkgyfddnycrmpdqrpmyeyctldphtrweycaikrcadnltmdvpl	300
QY	301	EPTTECLOGEGEYRGTVMVTIMNGILPCQRMDSQYRHEHDMTPENKCKDKLRENYCRNPD	360
Db	301	etteclqgggeyrgtvmvltimngilpcqrwdsqyphenmdtpefkckdlienyocrnpds	360
Db	361	ESPMCTTDPNIRVIGCSQIPNCDMSHGDCYRGNGKNYMGNTLSQTRSLTCSMDKNME	420
QY	421	DLHRHIFEMPRDASKINENYCRNPDDAHGPCYGNLFIWDCPISRCGGDTPTITVNL	480
Db	421	dlhrhlfemprdascklneynycrnpddahgpcygnlfiwdcplsrcegdtpitvnl	480
QY	481	DHPVISCATKOLRVNIGIPRTNIGMWNLSLRNKKICGSLIKESWVLTAROCPPSRD	540
Db	481	dhpviscatkqlrvnigiprtnlgmwnlslyrnklicgsliskewvltarqcpsrd	540
QY	541	LMDYENALGIDHYHNRGRGKCKOVNLNSQLYGREGSDLYMLKLRPAVLDFVSTIDLP	600
Db	541	lmdyenalgidhyhnrgrgkckvlnvslvygspgsdlvmlklarpavldfvsidlp	600
QY	601	NGSGTIPERTSCSYGMYGTGLINVDGLIRVAHLITMGNEKCSOHHRGKVTINSESLCAG	660
Db	601	ngsgtipektcsygywgtglindgllirvanhlylmgnkcsqhnrgvctlneseicag	660
QY	661	AEKIGSGPCEGDDGGLVCEQHKKRMVLTGVLPGCGCAINRGEIFVRAVYAKWIKIT	720
Db	661	aeklgsppcegdggglvceqhkrmvltgvlpgcgcaipnrgeifvrvayakwikit	720
QY	721	LTYKXVPOS 728	
Db	721	ltykvpos 728	
QY	4	AA559922	
Db	4	AA559922	
QY	21-DEC-1998	(first entry)	
Db	21-DEC-1998	(first entry)	
QY	Human leukocyte-derived hepatocyte growth factor HCL3.		
Db	Human leukocyte-derived hepatocyte growth factor HCL3.		
QY	Hepa2ocyte-derived growth factor; HGF; human; liver; hepatoma; diagnosis; therapy.		
Db	Hepa2ocyte-derived growth factor; HGF; human; liver; hepatoma; diagnosis; therapy.		
QY	Homo sapiens.		
Db	Homo sapiens.		
QY	EP859009-A2.		
Db	EP859009-A2.		
QY	19-AUG-1998.		
Db	19-AUG-1998.		
QY	07-JUN-1991.	91EP-0109369.	
Db	07-JUN-1991.	91EP-0109369.	
QY	11-JUN-1990.	90JP-0152474.	
Db	11-JUN-1990.	90JP-0152474.	

PA (NAKA/) NAKAMURA T.
XX
PI Asami O, Hagiya M, Ihara I, Nakamura T, Sakaguchi M,
PI Seki T, Shimizu S, Shimonishi M.
XX
XX
DR WPI: 1998-429650/37.
DR N-PSDB: AAV53626.
XX
XX
XX Preparation of recombinant hepatocyte growth factor polypeptide - by
PT culturing mammalian cells transformed with vector containing human
PT leukocyte-derived HGF gene
XX
XX
PS Claim 1, Fig 2a-c; 30pp: English.

Claim 1; Fig 2a-c; 30pp; English.

This is the amino acid sequence for human leukocyte-derived hepatocyte growth factor (HGF) HLC3, deduced from a cDNA clone (see AAW53656) obtained from a Leukocyte cDNA library. Another HGF, HLC2 (see AAW59293), was identified that differs from HLC3 by having amino acid residues 162-166 deleted. A claimed method of producing a HGF comprises: transforming mammalian cells (preferably CHO cells) with a recombinant expression vector that has a promoter (preferably the SV40 early promoter), a gene coding for human leukocyte-derived HGF, and the dihydrofolate reductase gene; (b) culturing the transformed cells in the presence of successively elevated concentrations of methotrexate; and (c) recovering the HGF polypeptide from the culture supernatant. HGF polypeptides may be used as hepatocyte culture reagents, liver regeneration promoters, in basic research on liver function, research on the action of various hormones and drugs on hepatocytes, research on the carcinogenesis mechanism of hepatoma, clinical diagnostic reagents using an antibody against the polypeptide and therapeutic drugs for liver disease.

Sequence 728 AA;

Query Match	100.0%;	Score 4126;	DB 19;	Length 728;
Best Local Similarity	100.0%;	Pred. No. 9.5e-277;		
Matches 728; Conservative	0;	Mismatches	0;	Gaps 0

OY	1	MMKKLLPALLHGVLLHLLHLLPLAAYAEQRRRRNTHEFFKSATYTLIKIDPALKIK	60
Db	1	mmvckllpalllqvvllhllllpilaayaeqrkrntllheffksatclliklidalik	60
OY	61	TKKVNATDOCANRCTRNKGLPFTCKAEVFDKARKOCMLPFNSKSSGVKKEEGHEPLYE	120
Db	61	tkkvntadgancrctrnkglpftckaelvfdkarkqclwfpnsmssgvkkekhegfhefplye	120
OY	121	NKQVIRNCIIGKGSYGYGYSITSSGIKQDPWSSMIFHEHSFLPSSYRGKDLOBNTCRNP	180
Db	121	nkqvlrnciilgkgsygygyvsltksgikqdpwssamlphensflpssyrgkdloqenyrnp	180
OY	181	RGEEGGPMCTSNSEVEYEVCDIIPQCESEVCMTCNGESYSGLMDDHTSSGKICQRMHIQTP	240
Db	181	rgeegsgpmctfnspevryevcdipqcevecmctngesyrglmddhesyglcqwdhqrtp	240
OY	241	HREKFLPERIPDKGFDDNYCRNPDGQRPWCYTLPDHTREYCAIKTCADNTMNDTVPL	300
Db	241	hrnkflperypdkgidnycrnpdgqrpwcylpdrhtreycaiktcadntmndtvpvl	300
OY	301	ETTECCIOGQEGRGYVNTLWNGIIPQGRMSQVPHENDMTPEMFCKCKDLRENTCRNPDS	360
Db	301	etteciaqggqegrygvntllwngiipqgrwdsqvpheindmtpemfckckdlrencycrnpdgs	360
OY	361	ESPMCFPTTDNIRVQYCSQIPNCDMSHQDCYRGNGKNGYMLNSQTRSGLTCSGMPKNNK	420
Db	361	espmcftfdpnirvqycsqipncdmshtgdcyrgngknygmnlsgtrsgltcsmwknnkme	420
OY	421	DLRRHIFWEDASKLMENTCRNPDDAHGWCYTGNPLIPWDYCPISRCBGDTPTIVNL	480
Db	421	dlrrhlfweddasklmenyrcnpddahgwcytgnpllpwdycpilsrcegdtptrivnl	480

QY 481 DHPVISC AKTKOLRVNNGIPTRTNIGMWSLRYRNKHIICGSLIKESVLTAROCFSPSRD 540
 |||||
 Db 481 dhpvisacktkqlrvnngiprtinigmwslryrnkhiicgslikesvltarqcfpsrd 540
 QY 541 LKDYEAWLGIDHVGKDECKOYLVNSQLYVGPBGSDVLMKLARPAVDDFVSTIDLP 600
 |||||
 Db 541 lkdyeawlgidhvgkdeckqylvnsqlyvgpbgdvlmklarparvlddfvstidlp 600
 QY 601 NYGCTIPEKTSVYGGVGYGLINYGDLIRVAHLXYIMGNEKCSOHHGKVTLNSEIICAG 660
 |||||
 Db 601 nygctipektsvgygyglinydglirvahlyimgnecsqnhrgkvlnseicag 660
 QY 661 AEKIGSGPCGEGDYGPLVCEQHKKRMVLYIVPGRGCAIPNRPGEFVRAVYAKWIHKII 720
 |||||
 Db 661 aekigsgpcegdyygplvceqghkrmvlyivpgrgcaipnrpgeifrvayakwiikhii 720
 QY 721 LTYKVPQS 728
 |||||
 Db 721 ltykvpqs 728

RESULT 5

AAW58696
 ID AAW58696 standard; protein; 728 AA.

XX AAW58696;
 AC

XX 08-SEP-1998 (first entry)
 DT

XX Human hepatocyte growth factor.
 DE

XX Human hepatocyte growth factor; HGF; high blood sugar.
 KM

XX Homo sapiens.
 OS

XX JP10167981-A.
 PN

XX 23-JUN-1998.
 PD

XX 09-DEC-1996; 96JP-0328357.
 PF

XX 09-DEC-1996; 96JP-0328357.
 PR

XX (MITU) MITSUBISHI CHEM CORP.
 PA

XX WPI; 1998-408607/35.
 DX

New preventive and/or treating agent - comprises hepatocyte growth factor, used for diseases caused by high blood sugar

XX Claim 5; Page 4-6; 6pp; Japanese.
 PS

XX The present sequence represents human hepatocyte growth factor (HGF).
 CC

XX A preventive and/or treating agent for the diseases caused by high
 CC blood sugar has been developed. The agent comprises HGF as an active
 CC component. The preventive and/or treating agent causes no excessively
 CC low blood sugar.

XX Sequence 728 AA;
 SQ

Query Match 100.0%; Score 4126; DB 19; Length 728;
 Best Local Similarity 100.0%; Pred. No. 9.5e-277;

Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWYTKLLPALLGHVLLHLLLPALPAIYAGQRRKNTIEFFKSAKTTLIKIDPALKIK 60
 |||||
 Db 1 mwvtkllpalllghvllhlllpalpaiaegqrkrntieffksaktlikidpalkik 60
 QY 61 TKKVNTADOCANRCTRKKGLPFTCKAFVFDKARKQCLMPFNSSSGVKKFEFGHEFDLYE 120
 |||||
 Db 61 tkkvntadgancrctrnkglpftckafvfdkarkqclwfpnsmssgvkfefghefdlye 120

QY 121 NKDYIRNCIIGKSGRYKGVSTITKGIKQFWSMIPHEHSFLPSSYRGKDIOENYCRNP 180
 |||||
 Db 121 nkdyirnciigksgrykgtvstitsgikqfwsmi p h e h s f l p s s y r g k d i o e n y c r n p 180
 QY 181 RGEEGGFWCFTSNPEVRYEVCDDIPQCSBVECMTCNGESYKGLMDHTSGKICQRMHQTP 240
 |||||
 Db 181 rgeeggfwcftsnpevryevcdipqcsvecmctcngesyrglmdhtsgkicqrmhqt p 240
 QY 241 HRHKEFLERYBDKFDNNYCNBNPDGPRPWCYTLDPHTRMRYCAIKKCADNTMNDTPVPL 300
 |||||
 Db 241 hrhkeflerybdkfdnnycnbnpdgprpwcycl d p h t r m r y c a i k k c a d n t m n d t p v p l 300
 QY 301 ETPECIOGCGEGYGTAVNTIWNIGIPCORWDSQYRPHEDMTPENKCKDLRENYCRNPDGS 360
 |||||
 Db 301 etpecio g c g e g y g t a v n t i w n i g i p c o r w d s q y r p h e d m t p e n k c k d l r e n y c r n p d g s 360
 QY 361 ESPWCFTTDPNIRVGYCSQIPNCDMSHGODCYRGNKKNYMGNSQTRSGLTCSKWDKNME 420
 |||||
 Db 361 espwcttdp n i r v g y c s q i p n c d m s h g o d c y r g n k k n y m g n s q t r s g l t c s k w d k n m e 420
 QY 421 DLHRHIFNEPDAKTLNENYCNBDDDAHGPWCYGNPLIPWDYCPISRCBGDTPTIVNL 480
 |||||
 Db 421 dlhrhifnepdasktlnenycnbdddahgpwcygnplipwdycp is r c b g d t p t i v n l 480
 QY 481 DHPVISC AKTKOLRVNNGIPTRTNIGMWSLRYRNKHIICGSLIKESVLTAROCFSPSRD 540
 |||||
 Db 481 dhpvisacktkqlrvnngiprtinigmwslryrnkhiicgslikesvltarqcfpsrd 540
 QY 541 LKDYEAWLGIDHVGKDECKOYLVNSQLYVGPBGSDVLMKLARPAVDDFVSTIDLP 600
 |||||
 Db 541 lkdyeawlgidhvgkdeckqylvnsqlyvgpbgdvlmklarparvlddfvstidlp 600
 QY 601 NYGCTIPEKTSVYGGVGYGLINYGDLIRVAHLXYIMGNEKCSOHHGKVTLNSEIICAG 660
 |||||
 Db 601 nygctipektsvgygyglinydglirvahlyimgnecsqnhrgkvlnseicag 660
 QY 661 AEKIGSGPCGEGDYGPLVCEQHKKRMVLYIVPGRGCAIPNRPGEFVRAVYAKWIHKII 720
 |||||
 Db 661 aekigsgpcegdyygplvceqghkrmvlyivpgrgcaipnrpgeifrvayakwiikhii 720
 QY 721 LTYKVPQS 728
 |||||
 Db 721 ltykvpqs 728

RESULT 6

AAW42998
 ID AAW42998 standard; protein; 728 AA.

XX AAW42998;
 AC

XX 23-JUL-1998 (first entry)
 DT

XX Recombinant human hepatocyte growth factor (HGF).
 DE

XX hepatocyte growth factor; HGF; W/O/W emulsion; blood half life;
 KM therapeutic; drug; hepatic tissue; hepatic disease; acute hepatitis;
 KW chronic hepatitis.

XX Homo sapiens.
 OS

XX JP10007587-A.
 PN

XX 13-JAN-1998.
 PD

XX 24-JUN-1996; 96JP-0163063.
 PF

XX 24-JUN-1996; 96JP-0163063.
 PR

XX (AZUM/) AZUMA H.
 PA

XX (MITU) MITSUBISHI CHEM CORP.
 PA

XX (MIYA-) MIYAZAKI KEN.
 PA

XX (TSUB/) TSUBOUCHI H.
 PA

XX WPI: 1998-126093/12.
 XX Preparation of emulsified composition - contains hepatocyte growth
 PT factor
 XX
 XX Claim 5: Pages 7-8; 9pp; Japanese.
 XX
 CC The present sequence represents a recombinant human hepatocyte growth
 CC factor (HGF). The protein is used in the inner aqueous phase of a
 CC W/O/W emulsion. The emulsion also contains an oil phase and an outer
 CC aqueous phase. The emulsified composition can considerably extend the
 CC blood half life of HGF. The emulsion is used in therapeutical drugs to
 CC deliver HGF to hepatic tissues in patients. It is useful for the
 CC treatment of hepatic diseases such as acute hepatitis and chronic
 CC hepatitis.
 CC
 XX
 XX Sequence 728 AA;

Query Match 100.0%; Score 4126; DB 19; Length 728;
 Best Local Similarity 100.0%; Pred. No. 9, 5e-277;
 Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTVKLLPALLOHVLHLLELLPIAIPYAEQORRRMTIHEFKSAKTTLIKIDPALKIK 60
 DB 1 mmtvkllpalllqhvllhlllplaiipyaegqrkrntlinhefkksaktllikidpalkik 60
 QY 61 TKKVNADQCANRCTRNKGLPFTCKAFVFDKARKQCLMFPPNSMSGYKKEFGHEFDLYE 120
 DB 61 tkkvnadqcanrctrnkglpftckafvfdkarkqclwfpnsmsgykkefghfedlye 120
 QY 121 NKQYIRNCITIGKRSYKGTVSITSGIKCOPWSSMIPHESEFLSSYRGKDLQENYCNRP 180
 DB 121 nkdyirncilgkrsygtvsitksikcqpssmiphefslssyrgkdlqenycnrcnp 180
 QY 181 RGEEGPWCFTSNPEVREYVCDIPQCSVEECMTNGESYRGIMDHTESGKICQRMHQTP 240
 DB 181 rgeegpwcftsnpevreyvcdipqcsveecmtngesyrgimhtesgkicqrmhqtpr 240
 QY 241 HRHKFLPERYDCKGFDNDYCNRPDQPRPWCYTLDPIHTRWECALKTCADMTMDTVPL 300
 DB 241 hrhkfleperrydckgfdndycnrcpdpwpwcytldphtreycaiktcadmtmdtdvpl 300
 QY 301 ETTTCIOGOGRGYGTNTWNGIIPCQRWDSQYRPHENDMPENFKCKDLRENYCNPDS 360
 DB 301 ettcioogogrgygtntwngiipcqrwdsqyrphendmpenfckckdlrenycnrdps 360
 QY 361 ESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRNGKNYMGNLQSTRSGLTCSMWDKNME 420
 DB 361 espwctftdpnirvgycsqipncdmshgdcyrgngknymgnlsqtrsgltcsmwdkme 420
 QY 421 DLNHRITWEPASKLNEYCNRPDDANGPWCYTGNPLIPMDYCRISCEGDTPTTYNL 480
 DB 421 dlhnhritwepasklneycnrcpddangpwcytgnplipmdycriscegdtptrtynl 480
 QY 481 DHPYISCKTQQLRVNIGIPRTNIGMWSLRYRNKNIICGSLIKESWYLAROCFSPSRD 540
 DB 481 dhpviscktkqlrvnigiprtnigmwslryrnknicgsslikeswylarocfspd 540
 QY 541 LKQYEAWLGIHDVHGRGDECKQVNLNSQVLYGREGSDVLMLKARPAVLDDFVSTIDLP 600
 DB 541 lkdyeawlgihdvhgrgdeckqylnvnsqlyvgregsdvlmlkarpvlddfvstidlp 600
 QY 601 NYGCTIEPKTCSYVGMWYTLINVDGLLRVAHLITMGENKCSQHNRRKVLINSEITAG 660
 DB 601 nygctiepktcsyvwmwytlinvdglrrvahlitmgenkcsqhnrrkvlineseitag 660
 QY 661 AEKIGSGCEGDYGPALVCEQHKMMWLVGVIYVPGRCALIPNRPGFVVAAYAKIHKII 720
 DB 661 aekigsgcedygpaltvceqhkmmwlvgivypgrcalipnrpgrfvvaayakihkii 720
 QY 721 LTYKVPQS 728

Db 721 ltykvpqs 728

RESULT 7

AAW39207
 ID AAW39207 standard; protein; 728 AA.

AC AAW39207;

DT 11-MAY-1998 (first entry)

XX Human hepatocyte growth factor protein.

DE Hepatocyte growth factor; HGF; parenchymatous cell proliferation factor;

KW insulin-like; human; preventative agent; therapeutic agent; glycogen;

KW hypoglycaemia.

OS Homo sapiens.

PN JP10007586-A.

PD 13-JAN-1998.

PF 05-JUN-1996; 96JP-0142839.

PR 26-APR-1996; 96JP-0108263.

PA (MITU) MITSUBISHI CHEM CORP.

DR WPI: 1998-133629/13.

XX Composition for treating conditions caused by insulin agonists -

PT e.g. hypoglycemia and glycogen diseases

PS Claim 5; Page 5-6; 8pp; Japanese.

XX This sequence represents a human hepatocyte growth factor (HGF), also

CC known as a hepat parenchymatous cell proliferation factor which can be

CC used as a preventative and/or therapeutic agent for diseases caused by

CC insulin-like action, e.g. hypoglycemia and glycogen disease. This factor

CC has activity to proliferate hepat parenchymatous cells and has a strong

CC affinity with heparin.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Db 301 ettecigqggyrgtvtlnimgipcqrwdsqyphndtptenfkckdlrenycrnpgds 360
QY 361 ESPWCFTTDPNIRVGYCSQIIPNCDMSHGDCYRGNGKNYMGNTLSQTRSGLTCSMMDK NME 420
Db 361 espwctftdpnirvgycsqipncdmshgdcyrngknymgnlsgtrsgltcsmmdk nme 420
QY 421 DLHRHIFWEPDASKLNNENYCRNPDDDAHGPWCYTGNPLIPWDYCPISRCGDTPTIVNL 480
Db 421 dlhrhifwepdascklneenycrnppddahgpwcytgnplipwdycpisrcegdtptrivnl 480
QY 481 DHPVISCAKTQLRVVNGIPTFRNIGMWVSLRRRNKHIGGSLIKESWVLTARQCPSRD 540
Db 481 dhpviscaaktqlrvvngiptfrnigmwslrrrnkhicgsslikesswvltarqcpsrd 540
QY 541 LKDYEAWLGIHDVHGRDECKQVNLVNSQLVYGPESDVLMLKARPVLDDEFTSTDLR 600
Db 541 lkdyeawlgihdvhgrdeckqvlvnsqlyvypesdvlmlkarpavlddftstcldr 600
QY 601 NYGCTIPEKTSQSVYGMGYTGLINVDGLLRVAHLTYMGNEKCSQHRGKVTLINESEICAG 660
Db 601 nygctipektsqsvygytglingdlrvahlyimgneksqhrgkvtlinesicag 660
QY 661 AAKIGSGPCEGDYGGPLVCEQHKRMVLYIVPGRGCAIPNRGIFVRAVYAKWIKIT 720
Db 661 aekigsqpcgedy99plvceqhkrmvlyivpgrgcaipnrp9lfrvayayakwihkii 720
QY 721 LTYKVPQS 728
Db 721 ltykvpqs 728

RESULT 8

AAR39521
ID AAR39521 standard; protein; 728 AA.

XX AAR39521:
AC 20-SEP-1993 (first entry)
DT 20-SEP-1993 (first entry)
XX Hepatocyte growth factor.
DE Hepatocyte growth factor.
KW HGF; side effect; cancer; tumour; chemotherapy; radiotherapy;
KW carcinostatic.
XX Synthetic.
OS WO9308821-A.
XX 13-MAY-1993.
XX 05-NOV-1992; 92MO-JP01433.
XX 07-NOV-1991; 91JP-0321412.
XX (NAKA/) NAKAMURA T.
XX Nakamura T.
XX WPI; 1993-167384/20.
XX N-PSDB; AAQ46040.
XX Side effect inhibitor for cancer therapy - containing hepatocyte
PT growth factor and inhibiting side effects resulting from
PT treatment by chemotherapy, radiotherapy
PS Disclosure; Fig 1; 36pp; Japanese.
XX Since HGF alleviates damage to normal cells and tissues, it can
XX inhibit side effects in treating cancers by chemotherapy,
XX radiotherapy, etc. With it, it is possible to conduct more intense
XX cancer therapy and to improve the carcinostatic effects. It is
XX extremely useful in the field of clinical medicine.

SQ Sequence 728 AA:

Query Match 99.9%; Score 4120; DB 14; Length 728;
Best Local Similarity 99.9%; Pred. No. 2,5e-276;
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MWYTKLIPALLQHVLLHLLELPALPYAEGORRRRTIEFFKSAKTLIKIDPALIK 60
Db 1 mwytklipalllqhvllhllelpalpyaeagqrkrntlieffksaktlikidpalkik 60
QY 61 TKRVNTADOCANBCTRNKGLPTCKAEVFDKARKQCLMPFPNSMSGVKKFGEHEDLYE 120
Db 61 tkrvntadocanbctrnkglptckaevfdkarkqclmfpnsmsgvkkfgehfedlye 120
QY 121 NKQYIRNCITGKRSYKGYSTIRKSGIKQCPNMSMTPHENSFLRPSYRGCDQENCRNP 180
Db 121 nkdyirncilgkrsygytstirksqikcpwssmlphensflpsyrghdigenycrnp 180
QY 181 RGEEGPWCFTSNPEVREVECDIPQCEVECMTCNGESYRGIMDHTESGKICQRMHQTP 240
Db 181 rgeegpwcftsnpevreyevcdipqcevecmtcngesyrgimdhthesgkicqrwdhqtpr 240
QY 241 HRHKFLPERYPDKGFDDNYCRNPDGQRPWCYTLDPTFRMEYCAIKTCADNTMNDVPL 300
Db 241 hrnkflperypdkgfdndnycrnpdgqrpwcytldpfrmeycalctcadntmndtvppl 300
QY 301 EETECTOGGEGYRGVTNTIWNGLPCQRMDSQYRHNEDMTPEMFKCKDLRENYCRNPDGS 360
Db 301 ettecigqggyrgtvtlnimgipcqrwdsqyphndtptenfkckdlrenycrnpgds 360
QY 361 ESPWCFTTDPNIRVGYCSQIIPNCDMSHGDCYRGNGKNYMGNTLSQTRSGLTCSMMDK NME 420
Db 361 espwctftdpnirvgycsqipncdmshgdcyrngknymgnlsgtrsgltcsmmdk nme 420
QY 421 DLHRHIFWEPDASKLNNENYCRNPDDDAHGPWCYTGNPLIPWDYCPISRCGDTPTIVNL 480
Db 421 dlhrhifwepdascklneenycrnppddahgpwcytgnplipwdycpisrcegdtptrivnl 480
QY 481 DHPVISCAKTQLRVVNGIPTFRNIGMWVSLRRRNKHIGGSLIKESWVLTARQCPSRD 540
Db 481 dhpviscaaktqlrvvngiptfrnigmwslrrrnkhicgsslikesswvltarqcpsrd 540
QY 541 LKDYEAWLGIHDVHGRDECKQVNLVNSQLVYGPESDVLMLKARPVLDDEFTSTDLR 600
Db 541 lkdyeawlgihdvhgrdeckqvlvnsqlyvypesdvlmlkarpavlddftstcldr 600
QY 601 NYGCTIPEKTSQSVYGMGYTGLINVDGLLRVAHLTYMGNEKCSQHRGKVTLINESEICAG 660
Db 601 nygctipektsqsvygytglingdlrvahlyimgneksqhrgkvtlinesicag 660
QY 661 AAKIGSGPCEGDYGGPLVCEQHKRMVLYIVPGRGCAIPNRGIFVRAVYAKWIKIT 720
Db 661 aekigsqpcgedy99plvceqhkrmvlyivpgrgcaipnrp9lfrvayayakwihkii 720
QY 721 LTYKVPQS 728
Db 721 ltykvpqs 728

RESULT 9

AAR25676
ID AAR25676 standard; protein; 728 AA.

XX AAR25676:
AC 20-JAN-1993 (first entry)
DT 20-JAN-1993 (first entry)
XX Recombinant human hepatocyte growth factor.
DE Recombinant human hepatocyte growth factor.
KW HGF; enhance growth; preparing transgenic animals; hepatic disease;
KW clinical diagnostic reagent; drug.
XX

Db	61	tkkvnltadqcanrctrnkqjrlpftckatvfdarkqclwfrfnsmssgyvkkneifhfdlye	120
Qy	121	NKDYIRNCLIGKGRSYKGTVSITRSGIKCPWSSMIEHESFLPSSYRGKDLQENYCRNP	180
Db	121	nkdylnrncilgkgrsykgtvstksqilkcqpwssmiphehsflpssyrgdldenyccrp	180
Qy	161	RGEEGPWCFTSNPEVREVCVDIPQCSEVBCMTGNSGTSGYGLMDHTEBSGKICORMDHQTP	240
Db	161	rgeegpwcftsnpevryevcdipqcsevecmctcnsgesyrjlmhntessgkicqrwdhqtp	240
Qy	241	HRHKLRPRYRDKGDDDDYCNBNRGORPWCYTLDPTHTHREYCAIKTCANTNNDDTVPL	300
Db	241	hnhkrlpryrdkgrdddnycnbnrgorpwcytldphtthreycalktcanntmddtvpl	300
Qy	301	ETTEECIOGSGYSGYGTVNTINNGIIPCCORMDSQYPHEDMTPENEKCKDLLENYCRNDGS	360
Db	301	etteeciqgsgygygtvntinngipcdqrwsdqyphedmtpeenekckdlleenyccrnpdgs	360
Db	361	ESPWCFTTDPNIRVGYCSQIRPNCDSHSGDCYRGNGKNTWGNLSQTRSGILTCSMDKNME	420
Qy	421	DLHRIIEHPRPASKLNEYCNBNRDDAHGPMCTYGNLIPWDYCPISIRCGSGTTPRTVNL	480
Db	421	dlnhriiehprpasklneycnbnrddahgpmctygnlfpwdycpilsircgsgttrptvnl	480
Qy	481	DHPVSAKCTKQRLRVANGIPRTNIGMNVSLRYRNKHKICGSGILKESMVLTAQCPSRD	540
Db	481	dhpvsaekctkqrlgvnyngiprttnlgnmvslryrnkhkicgsgilkesmvltaqcpsrd	540
Qy	541	LKDYEAWLGIHDVHGRGDEKCKQYLVNYSQLYVGEPSDLYLMLLARPVALDPEVSTIDLP	600
Db	541	lkydyaewlgihdvhnrgdekckqylvnysqlyvgepsdlylmlkarpavlddfvstidlp	600
Qy	601	NYGCTIPFKTSGSYGWCYTGLINVDGLLRVAHLTYMGNEKCSQHHGKTYLNESEICAG	660
Db	601	nygctipfktsasygwcytglinydgllrvahlylmgnekscsqhhgktylneiselcag	660
Qy	721	LTYKVPQS 728	720
Db	721	ltykvpqs 728	720
BULF 11			
AAR20005			
AC	AAR20005 standard; Protein; 728 AA.		
XX	AAR20005;		
DT	24-MAR-1992 (first entry)		
DE	Human hepatocyte growth factor.		
XX	HGF; kidney regeneration; nephritis.		
KM	Homo sapiens.		
OS	EPA62549-A.		
PN	27-DEC-1991.		
PD	18-JUN-1991; 91EP-0109923.		
XX	19-JUN-1990; 90JP-0158841.		
PR	(TOYM) TOYO ROSEKI KK.		
XX	Nakamura T;		
PI			

[illegible]

QY 721 LTYKVPQS 728
|||||
Db 721 ltykvpqs 728

RESULT 12

AAW8529

ID AAW8529 standard; protein: 728 AA.

AC AAW8529;

DT 02-MAR-1999 (first entry)

DE Human hepatocyte growth factor (HGF) variant 1.

KW Human; HGF; hepatocyte growth factor; variant; HGF receptor; medicine;
heparan sulphate proteoglycan; cancer.OS Homo sapiens.
Synthetic.

Key Location/Qualifiers

FT Misc-difference 73 /label= R73E
FT /note= "wild-type Arg is replaced with Glu"
FT Misc-difference 76 /label= R76E
FT /note= "wild-type Arg is replaced with Glu"

PN W09851798-A1.

XX 19-NOV-1998.

PF 07-MAY-1998; 98MO-GB01318.

PR 10-MAY-1997; 97GB-0009453.

PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.

PI Birchmeier W, Cheraid E, Hartmann G;

XX WPI: 1999-024213/02.

DR Hepatocyte growth factor variants - binding to hepatocyte growth
factor receptor, useful in medicine e.g. for cancer therapy

PS Claim 6; Page -: 75pp; English.

The present sequence represents a human hepatocyte growth factor (HGF) variant. The HGF variants of the invention bind to the HGF receptor, but are substantially unable to bind heparan sulphate proteoglycan. The HGF variants are useful in medicine e.g. variants having the same effects on target cells as wt-HGF can be administered to patients requiring HGF, while antagonistic variants can be administered to patients requiring HGF antagonists. They are especially useful for treating cancer. Some of the variants have a longer circulatory half-life in vivo and a greater mitogenic activity than wt-HGF in rats. They are therefore useful therapeutically as wt-HGF, but may have superior effects in vivo, e.g. give greater tissue penetration and ability to reach cell or tissue compartments.

CC Note: This sequence is not provided in the specification; it has been created by modifying the wild-type HGF sequence.

SQ Sequence 728 AA;

Query Match 99.8%; Score 4116; DB 20; Length 728;
Best local Similarity 99.7%; Pred. No. 4.7e-276;

Matches 726; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MWYTKLLPALLLQHVLLHLPLIAIPYAEQGRKRNTHERRKSKYLLIIDLAKTK 60
|||||

Db 1 mwytklpalllqhvllhllllpiaipyaeqgrkrntherrkkaaktlklidpalklk 60
QY 61 TKKVNADOCANFCTRNKGLPFTCAFAVPDKARKOCLWEPFNSMSGVKKKEGFEDYE 120
Db 61 tkkvnadqcanectenkb1pftckafvdkarkqclwrfnsmssgykkkegfefdye 120
QY 121 NKDYIRNCIIGKRSYKGVSTKSGIKQCPWSSMIRPHEHSLPSRYGKDLQENYCRNP 180
Db 121 nkdyirnciigkrsykvstsksgikcpwssmiphehslpsrygkdldgenycrnp 180
QY 181 RGEEGPWCFTSNPEVRYEVCIDIPQSEVECTNGESYRGLMPTESGKLCQSRMDQTP 240
Db 181 rgeegpwcftsnpevryevcidipqsevecmctngesyrlmhtesgkicqsrmdtqtp 240
QY 241 HRRKFLPERYPKGFDDWYCRNPQCPRPKCYTLDPTRWETCAIKTCADNTNDTDPVL 300
Db 241 hrkflperypkghfddwycrnpqcprrpkcytldptrweycaiktcadntndtdvpl 300
QY 301 ETTETCIGQEGYRGTVNTINMGIPQCRWDSQYPHEHDMTPENFKCKDLRENYCRNPDS 360
Db 301 ettecigqgegyrgtvntinmgipocrwdsqyphehdmtpenfckckdlrenycrnpds 360
QY 361 ESPMCFITDPIIRVGYCSQIRPNCMSHGQDCYKNGKNYWNLSQTRSGLTCSMDPKNME 420
Db 361 espmcfitdpirvgycsqirpncdmsghgdcyrgngknynwlnsqtrsgltcsmwdkme 420
QY 421 DLRRHIFWEPDASKLNNYCRNPDDDAHGPWCYTGPNFLIPWDYCPISRCGDTPTIVNL 480
Db 421 dlrrhifwepasklennycrnpdddahgpwcytgpnflipwdycpistrcgdtptivnl 480
QY 481 DHPVSCAKTKOLRVYNGIPTRNIGMWSLRKRNHICGSLIKESWLTAROCFPRSD 540
Db 481 dhpvscaktkqlrvyngiptrnigmwslrrnkhicgslikeswltarocfprsd 540
QY 541 LKDYEAWLGIHDVHGRDEKCOVLNWSQVYGPESDYLWKLAPAYLDQFVSTIDLP 600
Db 541 lkdyeawlgihdvhgrdekcvlwnwsqvygpesdylwklarpayldqfstidlp 600
QY 601 NYGCTIPEKTSQSVYWGVTGTLINYGDLRAVHLIYMGNEKSOHHRGKVTLNESIEICAG 660
Db 601 nygctipektsqsvywgvtgltinygdlravhlymgneksohhrgkvtlneseicag 660
QY 661 AEKIGSGPEGDYGCPLYCEOHKRMVGLVIVGRGCAIPNRPGLFVYAYIAKVIHKKII 720
Db 661 aekigsgpegdygsplyveqghkrmvlgvivgrgcaipnrpglfrvayyakvikhkii 720
QY 721 LTYKVPQS 728
Db 721 ltykvpqs 728

RESULT 13

AAR10656

ID AAR10656 standard; protein: 727 AA.

AC AAR10656;

DT 16-APR-1991 (first entry)

DE Hepatic parenchymal cell growth factor.

KW Hepatic parenchymal cell growth factor; cirrhosis.

XX Homo sapiens.

OS Homo sapiens.

FH Key

FT Peptide

FT Peptide

FT Peptide

PN EP412557-A.

Location/Qualifiers

/label= hGHF active fragment

/label= hGHF active fragment

/label= hGHF active fragment

PD 13-FEB-1991.
 XX
 PF 10-AUG-1990; 90EP-0115397.
 XX
 PR 27-JUL-1990; 90JP-0200898.
 PR 11-AUG-1989; 89JP-0209449.
 PR 03-APR-1990; 90JP-0088592.
 XX
 PA (MITU) MITSUBISHI KASEI CORP.
 PI Kitamura N, Miyazawa K, Datkubara Y, Tsoubouchi H, Naka D;
 PI Takahashi, Matsui R, Yoshiyama Y;
 DR WPI; 1991-045716/07.
 DR N-PSDB; AAQ10489.
 XX
 PT Hepatic parenchymal cell growth factor - produced by recombinant
 PT methods, useful for e.g. treating cirrhosis
 PT

Disclosure; fig 2; 46pp; English.

CC This hepatic growth factor and its active fragments (see feature
 CC table) are encoded by gene(s) contained within the expression plasmid,
 CC PKCRGF-2. This is used to transform host cells for recombinant
 CC expression of the human hepatic parenchymal cell growth factor (hHGF).
 CC hHGF is useful in the treatment of hepatic diseases, e.g. cirr-
 CC hosis. It is produced in large ams., stably using this method.
 CC
 XX
 SQ Sequence 727 AA;

Query Match 99.7%; Score 4115; DB 12; Length 727;
 Best Local Similarity 99.9%; Pred. No. 5.5e-276;
 Matches 726; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 WTKLRLPALLQHYLHLHLPLAIPYAEGRKRRNTIHEFKSAKTTLIKIDPALKIKT 61
 DB 1 WVKLIPALLIQLHYLHLHLPLAIPYAEGRKRRNTIHEFKSAKTTLIKIDPALKIKT 60
 QY 62 KKVNTAQCANCRCRRNNGLPPTCKAFVFDKARQCLMFPFMSGSYKKEGHEFDLYEN 121
 DB 61 KKVNTAQCANCRCRRNNGLPPTCKAFVFDKARQCLMFPFMSGSYKKEGHEFDLYEN 120
 QY 122 KDYIRNCIIGKRSYKSTVSTTKSGIKQPMWSMIPHSPLPSSYRGKLDQENVYCRNPR 181
 DB 121 KDYIRNCIIGKRSYKSTVSTTKSGIKQPMWSMIPHSPLPSSYRGKLDQENVYCRNPR 180
 QY 182 GEEGPMCFSTSNPEVRYEVDIPQSEVECMTCNGESYRGIMDHTESGKICQRMDDHOTPH 241
 DB 181 GEEGPMCFSTSNPEVRYEVDIPQSEVECMTCNGESYRGIMDHTESGKICQRMDDHOTPH 240
 QY 242 RHKFLPERYPDKGDDNYCRRPDCQPRPWCYTLDPHTRMEYCAIKTCADNMNTDVPLE 301
 DB 241 RHKFLPERYPDKGDDNYCRRPDCQPRPWCYTLDPHTRMEYCAIKTCADNMNTDVPLE 300
 QY 302 TTECTIOGEGYRGCTVNTYINGICQRMDSOYPREHMTPEMFCKDLRENYCRNPDSE 361
 DB 301 TTECTIOGEGYRGCTVNTYINGICQRMDSOYPREHMTPEMFCKDLRENYCRNPDSE 360
 QY 362 SPWCFETDPNIRVYGCQIPNCDMSHGQDCYRGNGKNYGMNLSQTRSLGTCGSMDDKNMED 421
 DB 361 SPWCFETDPNIRVYGCQIPNCDMSHGQDCYRGNGKNYGMNLSQTRSLGTCGSMDDKNMED 420
 QY 422 LHRITFEPDASKINENYCRNPDADAGPWCYTGNPILPWYCPISRCGEGTTPPIYVLD 481
 DB 421 LHRITFEPDASKINENYCRNPDADAGPWCYTGNPILPWYCPISRCGEGTTPPIYVLD 480
 QY 482 HPIVSCAKTKOLRVNIGPTFTNIGMMVSLRYRNKHICGSLIKESWVLTAROCFPPSDL 541
 DB 481 HPIVSCAKTKOLRVNIGPTFTNIGMMVSLRYRNKHICGSLIKESWVLTAROCFPPSDL 540
 QY 542 KDYEAMTGIHNVHGRGDEKCKQVLNVSQLVYGPESGDLVIMKLARPAVLDDFVSTIDLPN 601
 DB 542 KDYEAMTGIHNVHGRGDEKCKQVLNVSQLVYGPESGDLVIMKLARPAVLDDFVSTIDLPN 600

DB 541 kDyEaWlGIdhVhgrgdeKckQvLnvsqLvygpegSdlvImklarpavlddfvstIdlpn 600
 QY 602 YGCTIPKTSQSVYGMWYTGILINDGLLRVAHLKYMGNKCSQHHRGKVTLENESEICAGA 661
 DB 601 YGCTIPKTSQSVYGMWYTGILINDGLLRVAHLKYMGNKCSQHHRGKVTLENESEICAGA 660
 QY 662 EKISGPDCEGDYGPPLVCEQHKRMVNLGVIPVGRGCAIPNRPGLFVRVAYYAKMIHKIIL 721
 DB 661 EKISGPDCEGDYGPPLVCEQHKRMVNLGVIPVGRGCAIPNRPGLFVRVAYYAKMIHKIIL 720
 QY 722 TYKVPQS 728
 DB 721 tykvyps 727

RESULT 14

AAW88532 standard; protein: 728 AA.

AAW88532;

02-MAR-1999 (first entry)

Human hepatocyte growth factor (HGF) variant 4.

Human; HGF; hepatocyte growth factor; variant; HGF receptor; medicine;
 heparan sulphate proteoglycan; cancer.

Homo sapiens.
 OS
 Synthetic.

Key Location/Qualifiers

Misc-difference 91 /label= K91E /note= "wild-type Lys is replaced with Glu"

Misc-difference 93 /label= K93E /note= "wild-type Lys is replaced with Glu"

Misc-difference 94 /label= K94E /note= "wild-type Lys is replaced with Glu"

W09851798-A1.

19-NOV-1998.

07-MAY-1998; 98WO-G801318.

10-MAY-1997; 97GB-0009453.

(IMCR) IMPERIAL CANCER RES TECHNOLOGY.

Birchmeier W, Gherardi E, Hartmann G;

WPI; 1999-024213/02.

Hepatocyte growth factor variants - binding to hepatocyte growth

factor receptor, useful in medicine e.g. for cancer therapy

Disclosure; Page -: 75pp; English.

The present sequence represents a human hepatocyte growth factor (HGF) variant. The HGF variants of the invention bind to the HGF receptor, but are substantially unable to bind heparan sulphate proteoglycan. The HGF variants are useful in medicine e.g. variants having the same effects on target cells as wt-HGF can be administered to patients requiring HGF, while antagonistic variants can be administered to patients requiring HGF antagonists. They are especially useful for treating cancer. Some of the variants have a longer circulatory half-life in vivo and a greater mitogenic activity than wt-HGF in rats. They are therefore useful therapeutically as wt-HGF, but may have superior effects in vivo, e.g. give greater tissue penetration and ability to reach cell or tissue compartments.

CC Note: This sequence is not provided in the specification; it has been
 CC created by modifying the wild-type HGF sequence.

XX Sequence 728 AA;

Query Match 99.7%; Score 4113; DB 20; Length 728;
 Best Local Similarity 99.6%; Pred. No. 7.5e-276;
 Matches 725; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MMTVKLLPALLQHVHLHLLELLPAILPAIPYAEQGRKRRNTIHEFKSAKTTLIKIDPALKIK 60
DB 1 mmtvkl1palllqhvlhlhl1lllpailpyaeqgrkrntihfeksaktlikidpalkik 60
QY 61 TKKNTAOCANRCRNRKGLPPTCKAFVFDKARKOCLMFPNMSSSGKKEGFHFDLYE 120
DB 61 tkkntaocanrcnrnkglpptckafvfdkarkoclmfpnmsssgkkegfhefdlye 120
QY 121 NKDYIRNCITIGKRSYKGTVAITRSKICQPMSSMIPEHSFLPSSYRGKDLQENYCRNP 180
DB 121 nkdyirncitigkrsykgvtaitrskicqpmssmipehsflpssyrgkdlqenycrnp 180
QY 181 RGEEGGPMCFISNPEVRYEVCDDIPQCSVEECMTNGESYRGIMDHTESGKICQRMHDQTP 240
DB 181 rgeeggpfcfsnpevryevcdipqcsveecmtngesyrglmdhtesgkicqrwdhqt 240
QY 241 HRHFLPERYDCKGFDNDYCRNPDGQPRPWCYTLDPHTRMEYCAIKTCADNTMNDVPL 300
DB 241 hrhflperydkgfddnycrnpdgqprpwcyltdphtrmewcaiktcadntmndtvp 300
QY 301 ETTECIGOGEGYRGTVNTIWMGIPCCQWDSQYPHEHDMTEENFKCKDLRENYCRNPDGS 360
DB 301 ettecigogegyrgtvntiwmgipccqwdsqyphehdmteenfckckdlrenycrnpdgs 360
QY 361 ESPMCFITDPMIRVGYCSQIIPNCDSHGQDCYRGNGKRYMGNLSQTRSGILCSMDKXME 420
DB 361 espmcfitdpmirvgycsqiipncdshgdcyrgngkrymgnlsqtrsgilcsmdkxme 420
QY 421 DLHRIIFWEPDASKINENYCRNPDDAHGPMCYTGNPLIPMDYCPISRCGDTPTIYNL 480
DB 421 dlhriifwepdaskinenycrnpddahgpmcytgnplipmdycpiscgdtptiynl 480
QY 481 DHPVISCATKQRLRVNNGIPTRTNIGMNVSLRYRNKHIIGSGSLIKESWLTAROCFSPSRD 540
DB 481 dhpvyscatkqlrvnngiprttngmnlvslryrnkhiigsgslikeswltarocfspd 540
QY 541 LKDYEAMLGINDVHGRGDECKOVLNVSQLYGGPESDVLMLKLARPAVLDDFVSTIDLP 600
DB 541 lkdyeamlgindvhgrgdeckovlnvsqlyggpessdvlmlklarpavlddfvstidlp 600
QY 601 NYGCTIPKETSQSYVGMGYTGLINVDGLLVAHLVIMGENKSOHHRGKVTLENESEICAG 660
DB 601 nygctipektsqsyvgytgllnvdgllrvahlvimgenksohhrgekvtlneeseicag 660
QY 661 AEKIGSGPCESDYGRLVCEQHKKRMVNLGVTPRGCAIPNRPGLFVAVAYAKMIHKII 720
DB 661 aekigsgpcesdygrlvceqhkrmvnlgvtprgcaipnrpglfvavayakmihkii 720
QY 721 LTYKVPQS 728
DB 721 ltykvpgs 728

```

RESULT 13

ID AARI5623 standard; Protein: 728 AA.

XX AARI5623:

DT 18-MAR-1992 (first entry)

XX Human leukocyte-derived HGF encoded by clone HLC2.

KW Hepatocyte growth factor; liver; hepatoma.

OS Homo sapiens.

XX EP461560-A.

XX 18-DEC-1991.

XX 07-JUN-1991; 91EP-0109369.

XX 11-JUN-1990; 90JP-0152474.

XX (TOYM) TOYO BOSEKI KK.

XX Nakamura T, Hagiya M, Seki T, Shimonishi M, Shimizu S;

XX Iherai, Sakaguchi M, Asami O;

XX WPI: 1991-370578/51.

XX N-PSDB: AAQ15176.

PT Recombinant human leukocyte-derived hepatocyte growth factor -

PT with DNA encoding it, recombinant expression vectors and

PT transformant cells expressing it.

CC Claim 2; Fig 2; 33pp: English.

CC The sequence was deduced from a portion of HLC2, one of two clones,

CC (for HLC3 see AARI5624) isolated from a cDNA library prepd. from mrNA

CC extracted from human leukocytes. HLC2 has almost the same sequence

CC as HLC3 except for five residues, 162-166, which do not appear in

CC HLC3. The DNA sequence can be expressed and the resulting protein,

CC recombinant HGF, used in hepatocyte cultivation, liver regenera-

CC tion, hepatocyte research, esp. the mechanism of hepatoma, and to

CC prepare anti-HGF antibodies for diagnosis and therapy.

XX Sequence 728 AA;

Query Match 99.7%; Score 4112; DB 12; Length 728;
 Best Local Similarity 99.7%; Pred. No. 8.8e-276;
 Matches 726; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MMTVKLLPALLQHVHLHLLELLPAILPAIPYAEQGRKRRNTIHEFKSAKTTLIKIDPALKIK 60
DB 1 mmtvkl1palllqhvlhlhl1lllpailpyaeqgrkrntihfeksaktlikidpalkik 60
QY 61 TKKNTAOCANRCRNRKGLPPTCKAFVFDKARKOCLMFPNMSSSGKKEGFHFDLYE 120
DB 61 tkkntaocanrcnrnkglpptckafvfdkarkoclmfpnmsssgkkegfhefdlye 120
QY 121 NKDYIRNCITIGKRSYKGTVAITRSKICQPMSSMIPEHSFLPSSYRGKDLQENYCRNP 180
DB 121 nkdyirncitigkrsykgvtaitrskicqpmssmipehsflpssyrgkdlqenycrnp 180
QY 181 RGEEGGPMCFISNPEVRYEVCDDIPQCSVEECMTNGESYRGIMDHTESGKICQRMHDQTP 240
DB 181 rgeeggpfcfsnpevryevcdipqcsveecmtngesyrglmdhtesgkicqrwdhqt 240
QY 241 HRHFLPERYDCKGFDNDYCRNPDGQPRPWCYTLDPHTRMEYCAIKTCADNTMNDVPL 300
DB 241 hrhflperydkgfddnycrnpdgqprpwcyltdphtrmewcaiktcadntmndtvp 300
QY 301 ETTECIGOGEGYRGTVNTIWMGIPCCQWDSQYPHEHDMTEENFKCKDLRENYCRNPDGS 360
DB 301 ettecigogegyrgtvntiwmgipccqwdsqyphehdmteenfckckdlrenycrnpdgs 360
QY 361 ESPMCFITDPMIRVGYCSQIIPNCDSHGQDCYRGNGKRYMGNLSQTRSGILCSMDKXME 420
DB 361 espmcfitdpmirvgycsqiipncdshgdcyrgngkrymgnlsqtrsgilcsmdkxme 420
QY 421 DLHRIIFWEPDASKINENYCRNPDDAHGPMCYTGNPLIPMDYCPISRCGDTPTIYNL 480
DB 421 dlhriifwepdaskinenycrnpddahgpmcytgnplipmdycpiscgdtptiynl 480

```

```
OY 481 DHPVISC AKTOLRVVNGIPTRTNIGMWSLARYRNKHICGSLIKESWLTAROCFPRS D 540
    |||||||
Db 481 dhpvisca ktqlrvvngl p l t n i g m w s l a r y r n k h i c g s l i k e s w l t a r o c f p r s d 540
OY 541 LKDYEA MLGIHDYHGRGDEKCKQVLNVSQLVGPESGDLVLMKLRAPVLDDEVSTIDL P 600
    |||||||
Db 541 lkdyeaw l g i h d v h r r g d e k c k q v l n v s q l v y g p e s g d l v l m k l a r p a v l d d e v s t i d l p 600
OY 601 NYCCTIPEKTSCSVYGMGTGGLINVDGLLRVAHLIYMGNEKCSQHHRGKVTLNSEI CAG 660
    |||||||
Db 601 nygcti p e k t s c s v y g m g t g g l i n v d g l l r v a h l i y m g n e k c s q h r g k v t l n s e i c a g 660
OY 661 AEKIGSGPCGEGDYGPVCEQHQHMRWLVGIVPGRGCAIPNRPGEFVRVAVYAKWIKH I I 720
    |||||||
Db 661 aekigs g p c e g d y g p l v c e q h k m r m v l g v l v p p r g c a l p n r p g e f v r v a y a k w i k h i l 720
OY 721 LTYKVPQS 728
    |||||||
    721 ltykvpqs 728
```

Search completed: June 18, 2002, 17:46:59
Job time: 240 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2002, 17:45:19 ; Search time 24.39 Seconds

(without alignments)
2868.103 Million cell updates/sec

Title: US-09-423-516-2

Perfect score: 4126
Sequence: 1 MWVTKLPALLQHVHLHL.....VAYYAKWIKIITLYKVPQS 728

Scoring table: BIOSUM62
Gapop 10.0 , Gapept 0.5

Indexed: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4126	100.0	728	1 JH0579	hepatocyte growth
2	3804	92.2	728	1 A60185	hepatocyte growth
3	3802	92.1	728	1 A35644	hepatocyte growth
4	2806	68.0	710	1 I51283	hepatocyte growth
5	1821.5	44.1	411	2 I51285	hepatocyte growth
6	1652	40.0	711	1 A47136	macrophage-stimula
7	1615	39.1	716	1 JCS061	macrophage-stimula
8	1608	39.0	716	1 A40332	macrophage-stimula
9	1446.5	35.1	790	1 PLPG	plasma (EC 3.4.21
10	1408.5	34.1	812	1 PLBO	plasma (EC 3.4.21
11	1379.5	33.4	812	1 PLMS	plasma (EC 3.4.21
12	1371	33.2	810	1 PLHU	plasma (EC 3.4.21
13	1369	33.2	810	2 B30848	plasma (EC 3.4.21
14	1353.5	32.8	810	2 I46260	plasma (EC 3.4.21
15	1212.5	29.4	4548	1 S00657	apoptin (EC
16	1188.5	28.8	1420	2 A32869	apoptin (EC
17	945	22.9	2869	2 T18518	apoptin (EC
18	818	19.8	460	2 B61545	plasma (EC 3.4.21
19	784.5	19.0	455	2 A61545	plasma (EC 3.4.21
20	559.5	13.6	559	1 A35029	t-plasminogen acti
21	551.5	13.4	559	1 A29941	t-plasminogen acti
22	543	13.2	625	1 TBBO	t-plasminogen acti
23	532.5	12.9	562	1 UKHUT	t-plasminogen acti
24	531.5	12.9	618	2 A35827	t-plasminogen acti
25	527.5	12.8	622	1 TBHU	thrombin (EC 3.4.2
26	518	12.6	617	2 S10511	thrombin (EC 3.4.2
27	510.5	12.4	603	2 S28941	thrombin (EC 3.4.2
28	480.5	11.6	655	1 A4688	coagulation factor
29	450	10.9	558	2 JCS878	hepatocyte growth

30	449	10.9	615	1 KFH012	coagulation factor
31	445	10.8	336	2 S33879	plasmin precursor
32	426.5	10.3	593	2 S45281	coagulation factor
33	423.5	10.3	560	1 JCA4795	plasma hyaluronan-
34	393	9.5	433	1 JN0560	u-plasminogen acti
35	390	9.5	431	1 UKHU	u-plasminogen acti
36	388	9.4	433	1 UKMS	u-plasminogen acti
37	386	9.4	442	1 UKPG	u-plasminogen acti
38	385	9.3	434	1 A35005	u-plasminogen acti
39	382.5	9.3	432	1 S18932	u-plasminogen acti
40	382.5	9.3	433	1 UKBAV	u-plasminogen acti
41	376.5	9.1	477	1 A43369	t-plasminogen acti
42	375	9.1	1035	1 A43090	enteropeptidase (E
43	373.5	9.1	477	2 JS0598	t-plasminogen acti
44	373	9.0	477	2 JS0597	t-plasminogen acti
45	372	9.0	273	2 A47246	trypsin (EC 3.4.2

ALIGNMENTS

RESULT 1

JH0579
hepatocyte growth factor precursor [validated] - human
N:Alternate names: hepatopoietin A; scatter factor
C:Species: Homo sapiens (man)
C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Dec-2000
C:Accession: JH0579; J00333; A41140; B36677; A36677; A33512; A39006; PH0114; A37796;
R:Seiki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
Gene 102, 213-219, 1991
A:Title: Organization of the human hepatocyte growth factor-encoding gene.
A:Reference number: JH0579; MUID:91340155
A:Accession: JH0579
A:Molecule type: DNA
A:Residues: 1-728 <SEK>
A:Cross-references: DDBJ:D90318
A:Note: the authors translated the codon GAA for residue 662 as Gly
R:Seiki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
Submitted to JIPID, March 1991
A:Description: Organization of the human hepatocyte growth factor-encoding gene.
A:Reference number: J00333
A:Accession: J00333
A:Molecule type: DNA
A:Residues: 1-481, 'RT', 484-728 <SE2>
R:Weidner, K.M.; Arakaki, N.; Hartman, G.; Vandekerckhove, J.; Weingart, S.; Rieder, Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991
A:Title: Evidence for the identity of human scatter factor and human hepatocyte growth factor.
A:Reference number: A41140; MUID:91334393
A:Accession: A41140
A:Molecule type: mRNA
A:Residues: 1-728 <ME1>
A:Cross-references: GB:M73239; NID:9337935; PIDN:AAA64239.1; PID:9337936
R:Seiki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya Biochem. Biophys. Res. Commun. 172, 321-327, 1990
A:Title: Isolation and expression of cDNA for different forms of hepatocyte growth fa
A:Reference number: A36677; MUID:91025062
A:Accession: B36677
A:Molecule type: mRNA
A:Residues: 1-728 <SE3>
A:Cross-references: GB:M60718; NID:9184031; PIDN:AAA52648.1; PID:9184032
A:Accession: A36677
A:Molecule type: mRNA
A:Residues: 1-161, 167-728 <SE4>
A:Cross-references: EMBL:X16322
A:Experimental source: Leukocyte
R: Miyazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nak Biochem. Biophys. Res. Commun. 163, 967-973, 1989
A:Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth
A:Reference number: A33512; MUID:89392017
A:Accession: A33512
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-728 <MIY>

A:Cross-references: GB:M29145; NID:g184041; PIDN:AA52650.1; PID:g306846
R:Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hilt
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
A:Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocyte
A:Reference number: A39006; MUID:9110540
A:Accession: A39006
A:Molecule type: mRNA
A:Residues: 1-161,167-728 <RUB>
A:Cross-references: GB:M55379
A:Experimental source: embryonic lung
R:Toshiyama, Y.; Araiaki, N.; Naka, D.; Takahashi, K.; Hirose, S.; Kondo, J.; Nakayama,
Biochem. Biophys. Res. Commun. 175, 660-667, 1991
A:Title: Identification of the N-terminal residue of the heavy chain of both native and
A:Reference number: PH0114; MUID:91207365
A:Accession: PH0114
A:Molecule type: protein
A:Residues: 32-43;53-58 <YOS>
A:Experimental source: plasma
R:Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.
J. Cell Biol. 111, 2097-2108, 1990
A:Title: Scatter factor: molecular characteristics and effect on the invasiveness of ep
A:Reference number: A37796; MUID:91035621
A:Accession: A37796
A:Molecule type: protein
A:Residues: 86-91;329-344;356-363, 'XX', 366-370;425-434;442-447, 'X', 449-450;543-546, 'X', 5
R:Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimonishi, M.; Sugimura, A.; Tash
Nature 342, 440-443, 1989
A:Title: Molecular cloning and expression of human hepatocyte growth factor.
A:Reference number: 506794; MUID:90066676
A:Accession: 506794
A:Molecule type: mRNA
A:Residues: 1-31, 'HK', 34-77, 'N', 79-292, 'V', 294-299, 'M', 301-316, 'A', 318-335, 'K', 337-386, '
A:Cross-references: EMBL:X16323; NID:g32081; PIDN:CAA34387.1; PID:g32082
A:Experimental source: liver
A:Note: the authors translated the codon CAG for residue 727 as Glu
A:Note: part of this sequence, including the amino end of both the alpha and beta chains
R:Hartman, G.; Naldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchm
Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth fact
A:Reference number: 159214; MUID:93087571
A:Accession: 159214
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-288, 'ET' <HAR>
A:Cross-references: GB:L02931; NID:g184033; PIDN:AA52649.1; PID:g184034
R:Myazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
Eur. J. Biochem. 197, 15-22, 1991
A:Title: An alternatively processed mRNA generated from human hepatocyte growth factor g
A:Reference number: S15443; MUID:91200041
A:Accession: S15443
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-288, 'ET' <MY2>
A:Cross-references: EMBL:X57574; NID:g32083; PIDN:CAA0802.1; PID:g32084
R:Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.
Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991
A:Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: cloning
A:Reference number: 152253; MUID:92062058
A:Accession: 152253
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 161-166 <SHI>
A:Cross-references: GB:S62561; NID:g237996; PIDN:AA520169.1; PID:g237997
C:Genetics:
A:Gene: GDB:HGF
A:Cross-references: GDB:127524; OMIM:142409
A:Map position: 7q21.1-7q21.1
A:Intons: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 469/1; 48
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Function:
A:Description: stimulates mitosis of hepatocytes and other cells
A:Note: does not have proteinase activity
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology

C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle;
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-43,495-728/Product: hepatocyte growth factor #status experimental <MAT>
F:32-43/Domain: alpha chain #status experimental <ACH>
F:32-43/Domain: alpha chain #status experimental <ACH>
F:128-206/Domain: kringle homology <KR1>
F:211-288/Domain: kringle homology <KR2>
F:305-383/Domain: kringle homology <KR3>
F:391-469/Domain: kringle homology <KR4>
F:495-728/Domain: beta chain #status experimental <BCH>
F:495-716/Domain: trypsin homology <TRY>
F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F:294,402,566,653/Binding site: carbonylde (Asn) (covalent) #status predicted
F:487-604/Disulfide bonds: #status predicted

Query Match 100.0%; Score 4126; DB 1; Length 728;
Best Local Similarity 100.0%; Pred. No. 2,4e-282;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MMTKLLPALLQVLLHLPLAIPAEQRRRRNTHEFKSAKTLIKIDPAKIK 60
1 MMTKLLPALLQVLLHLPLAIPAEQRRRRNTHEFKSAKTLIKIDPAKIK 60
Db 1 MMTKLLPALLQVLLHLPLAIPAEQRRRRNTHEFKSAKTLIKIDPAKIK 60
OY 61 TKKVTADQCANRCKTNGKLPFTCKAFVDFKARQCLMFPPNSSGYKKEGHEFDLYE 120
61 TKKVTADQCANRCKTNGKLPFTCKAFVDFKARQCLMFPPNSSGYKKEGHEFDLYE 120
Db 61 TKKVTADQCANRCKTNGKLPFTCKAFVDFKARQCLMFPPNSSGYKKEGHEFDLYE 120
OY 121 NKDYIRNCTIIGKRSYKTSVITKSGIKCOPWSSMIPHEHFLSSYKGLQENYCNP 180
121 NKDYIRNCTIIGKRSYKTSVITKSGIKCOPWSSMIPHEHFLSSYKGLQENYCNP 180
Db 121 NKDYIRNCTIIGKRSYKTSVITKSGIKCOPWSSMIPHEHFLSSYKGLQENYCNP 180
OY 181 RGEEGPWCFTSNPEVREYVCDIPQCEVEECMTGNGESYRGIMDHTEGKICQRMHDQTP 240
181 RGEEGPWCFTSNPEVREYVCDIPQCEVEECMTGNGESYRGIMDHTEGKICQRMHDQTP 240
Db 181 RGEEGPWCFTSNPEVREYVCDIPQCEVEECMTGNGESYRGIMDHTEGKICQRMHDQTP 240
OY 241 HRHKLPRRYDCKGFDNDYCNRPDQPPKCYTLDPHTRMEYCAIKTCADTMDTVPPL 300
241 HRHKLPRRYDCKGFDNDYCNRPDQPPKCYTLDPHTRMEYCAIKTCADTMDTVPPL 300
Db 241 HRHKLPRRYDCKGFDNDYCNRPDQPPKCYTLDPHTRMEYCAIKTCADTMDTVPPL 300
OY 301 ETTETIOGQGGYRGVTWITNGGIPQCRWDQSPHEHMTPEHFKCKDLRENYCRNPDGS 360
301 ETTETIOGQGGYRGVTWITNGGIPQCRWDQSPHEHMTPEHFKCKDLRENYCRNPDGS 360
Db 301 ETTETIOGQGGYRGVTWITNGGIPQCRWDQSPHEHMTPEHFKCKDLRENYCRNPDGS 360
OY 361 ESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNSLQSTRSGLTCSMWDKME 420
361 ESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNSLQSTRSGLTCSMWDKME 420
Db 361 ESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNSLQSTRSGLTCSMWDKME 420
OY 421 DLHRIIFEPDPAKINENYCNRPDDAHGPMCYTGNPLIPMDYCPISRCBEDTPTTYNL 480
421 DLHRIIFEPDPAKINENYCNRPDDAHGPMCYTGNPLIPMDYCPISRCBEDTPTTYNL 480
Db 421 DLHRIIFEPDPAKINENYCNRPDDAHGPMCYTGNPLIPMDYCPISRCBEDTPTTYNL 480
OY 481 DHPVTSCKATKQLRVNGIPRTNIGMWSLRYRNKHKICGSLIKESVNLTAAROCFSPSRD 540
481 DHPVTSCKATKQLRVNGIPRTNIGMWSLRYRNKHKICGSLIKESVNLTAAROCFSPSRD 540
Db 481 DHPVTSCKATKQLRVNGIPRTNIGMWSLRYRNKHKICGSLIKESVNLTAAROCFSPSRD 540
OY 541 LKDYEAWLGIHDVHGRGDEKCKOVLNSQVLYGREGSDLVLMKLARPAVLDPEVSTDLPL 600
541 LKDYEAWLGIHDVHGRGDEKCKOVLNSQVLYGREGSDLVLMKLARPAVLDPEVSTDLPL 600
Db 541 LKDYEAWLGIHDVHGRGDEKCKOVLNSQVLYGREGSDLVLMKLARPAVLDPEVSTDLPL 600
OY 601 NYGCTIPKRTKSCSYVGMKGTGLINVDGLLRVAHLTYMNEKCSQHHREKVTLNSEICAG 660
601 NYGCTIPKRTKSCSYVGMKGTGLINVDGLLRVAHLTYMNEKCSQHHREKVTLNSEICAG 660
Db 601 NYGCTIPKRTKSCSYVGMKGTGLINVDGLLRVAHLTYMNEKCSQHHREKVTLNSEICAG 660
OY 661 AEKISGCEGEGYGLVCEQHKMAMVGLVYPRGCAIPNRPGLFVAVATYAKWIKII 720
661 AEKISGCEGEGYGLVCEQHKMAMVGLVYPRGCAIPNRPGLFVAVATYAKWIKII 720
Db 661 AEKISGCEGEGYGLVCEQHKMAMVGLVYPRGCAIPNRPGLFVAVATYAKWIKII 720
OY 721 LTYKVPQS 728
721 LTYKVPQS 728
Db 721 LTYKVPQS 728

RESULT 2

A60185
hepatocyte growth factor precursor - mouse
N:Alternate names: hepatoleitin A: scatter factor
C:Species: Mus musculus (house mouse)
C:Date: 03-Mar-1993 #sequence-revision 26-May-1994 #text-change 16-Jun-2000
C:Accession: J02117; PC2064; A60185; S43416; S45521; S17173; S10966; I48758; J00231
R:Sasaki, M.; Nishio, M.; Sasaki, T.; Enami, J.
Biochem. Biophys. Res. Commun. 199, 772-779, 1994
A:Title: Identification of mouse mammary fibroblast-derived mammary growth factor as hep
A:Reference number: J02117; M01D:94183257
A:Accession: J02117
A:Molecule type: mRNA
A:Residues: 1-728 <SA2>
A:Cross-references: GB:D10212; NID:q220435; P1DN:BA01064.1; PID:q220436
A:Experimental source: fibroblast, COS-1 cell
A:Note: submitted to JIPID, May 1993
A:Accession: PC2064
A:Molecule type: protein
A:Residues: 496-504 <SA2>
Proc. Soc. Exp. Biol. Med. 195, 34-43, 1990
A:Title: Purified scatter factor stimulates epithelial and vascular endothelial cell mig
A:Reference number: A60185; M01D:90377927
A:Accession: A60185
A:Molecule type: protein
A:Residues: 'X',184-188,'KX',191-192,'X',194,'XX',197,357-364,'XX',367,375-377,'E',379,
R:Jiliu, Y.; Michalopoulos, G. K.; Zarnegar, R.
Biochim. Biophys. Acta 1216, 299-303, 1993
A:Title: Molecular cloning and characterization of cDNA encoding mouse hepatocyte growth
A:Reference number: S43416; M01D:94060105
A:Accession: S43416
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-728 <L1U>
A:Cross-references: EMBL:X72307
R:Jiliu, Y.
submitted to the EMBL Data Library, May 1993
A:Reference number: S45521
A:Accession: S45521
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-563,'H',565-728 <L12>
A:Cross-references: EMBL:X72307
R:Coffier, A.; Fellows, J.; Young, S.; Pappin, D.; Rahman, D.
Biochem. J. 278, 35-41, 1991
A:Title: Purification and characterization of biologically active scatter factor from ra
A:Reference number: S17173; M01D:91354223
A:Accession: S17173
A:Molecule type: protein
A:Residues: 496-517,'T',519 <COF>
R:Gherardi, E.; Stoker, M.
Nature 346, 228, 1990
A:Title: Hepatocytes and scatter factor.
A:Reference number: S10966; M01D:90326152
A:Accession: S10966
A>Status: preliminary
A:Molecule type: protein
A:Residues: 496-507,'X',509-512,'L',514-516,'X',518-519 <NAT>
R:Plaschke-Schluter, A.; Behrens, J.; Gherardi, E.; Birchmeier, W.
J. Biol. Chem. 270, 830-836, 1995
A:Title: Characterization of the scatter factor/hepatocyte growth factor gene promoter.
A:Reference number: I48758; M01D:95122532
A:Accession: I48758
A>Status: preliminary
A:Molecule type: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-30 <RES>
A:Cross-references: EMBL:X81630; NID:q673451; P1DN:CA57286.1; PID:q673452
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Function:
A:Description: stimulates mitosis of hepatocytes and other cells
A:Note: does not have proteinase activity
C:Superfamily: hepatocyte growth factor; kringe homology; trypsin homology
C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringe; pyz

F:1-32/Domain: signal sequence #status predicted <SIG>
F:56-495/Domain: hepatocyte growth factor #status predicted <MAT>
F:56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F:129-207/Domain: kringe homology <KR1>
F:212-289/Domain: kringe homology <KR2>
F:306-384/Domain: kringe homology <KR3>
F:392-470/Domain: kringe homology <KR4>
F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F:496-719/Domain: trypsin homology <TRY>
F:733/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F:295,403,569,656/Binding site: carbonyldate (asn) (covalent) #status predicted
F:488-607/Disulfide bonds: #status predicted

Query Match 92.2% Score 3804; DB 1; Length 728;
Best Local Similarity 90.8% Pred. No. 1e-259;
Matches 660; Conservative 39; Mismatches 26; Indels 2; Gaps 1;

QY 1 MMTVRLPALLOHVLHLHLPLAIPAEQRRRRRTHEFKSAKTLIKIDPALKIK 60
DB 2 MMTGRLPLVLLQHLHLHLHLHLHLAIPAEQKKRRRTLHEFKSAKTLIKEDPLKIK 61
QY 61 TKKYNFADQCANFTTRNKGDLPTCKAFVFDKARRQCLMFPNMSGGVKKFGEHFDLYE 120
DB 62 TKKNSADECANFTTRNKGDLPTCKAFVFDKARRQCLMFPNMSGGVKKFGEHFDLYE 121
QY 121 NKQVIRNCTIKGRSYKTVSITSGIKCQWSSMIPHEHFLSSYKGLQENYCNRP 180
DB 122 NKQVIRNCTIKGRSYKTVSITSGIKCQWSSMIPHEHFLSSYKGLQENYCNRP 181
QY 181 RGEFGPWCFTSNPEVREVEDIPQCEVEECMTNGESYRCLMDHTESGICQWMDHOTP 240
DB 182 RGEFGPWCFTSNPEVREVEDIPQCEVEECMTNGESYRCLMDHTESGICQWMDHOTP 241
QY 241 HRHKLPERYPDKGFDNDYCNPDGQPRPWCYTLDPTRWECYAIKTCADNTMNDTVPL 300
DB 242 HRHKLPERYPDKGFDNDYCNPDGQPRPWCYTLDPTRWECYAIKTCADNTMNDTVPL 301
QY 301 ETTECIGOGSGRYGTVTINNGIPIQCRWDSQYHEHDMTEENKCKDLRENYCNRPDS 360
DB 302 ETTECIGOGSGRYGTVTINNGIPIQCRWDSQYHEHDMTEENKCKDLRENYCNRPDS 361
QY 361 ESPMCFETDPNIRVGYCSQIIPNCDSHGQDYPKNGKMYMGNLSQTRSGILCSMMDKME 420
DB 362 ESPMCFETDPNIRVGYCSQIIPNCDSHGQDYPKNGKMYMGNLSQTRSGILCSMMDKME 421
QY 421 DLHRIHFWEPDASKLNNYCNPDGQPRPWCYTLDPTRWECYAIKTCADNTMNDTVPL 480
DB 422 DLHRIHFWEPDASKLNNYCNPDGQPRPWCYTLDPTRWECYAIKTCADNTMNDTVPL 481
QY 481 DHPVISCATKQLRVNGIPIRTNIGMNVSLRYRNKHTCGSSLIKESNVLTAROCFPRR- 539
DB 482 DHPVISCATKQLRVNGIPIRTNIGMNVSLRYRNKHTCGSSLIKESNVLTAROCFPRR- 541
QY 540 -DLKDYEAMLGIIHNVHGGDECKQOVNLVLYGPEGSDDLVLKMLARPAVLDFVSTID 598
DB 542 KDLKDYEAMLGIIHNVHGGDECKQOVNLVLYGPEGSDDLVLKMLARPAVLDFVSTID 601
QY 599 LPNVCCTIPEKTSYGVGWTGLINVDGLLRVAHLTYMGNEKSSQHHRGVTLNSESIC 658
DB 602 LPNVCCTIPEKTSYGVGWTGLINVDGLLRVAHLTYMGNEKSSQHHRGVTLNSESIC 661
QY 659 AGAERIGSGPEGSYGVGWTGLINVDGLLRVAHLTYMGNEKSSQHHRGVTLNSESIC 718
DB 662 AGAERIGSGPEGSYGVGWTGLINVDGLLRVAHLTYMGNEKSSQHHRGVTLNSESIC 721
QY 719 IILTYKV 725
DB 722 VILTYKL 728

RESULT 3
A35644

hepatocyte growth factor precursor - rat
 N:Alternate names: hepatopietin A; scatter factor
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 21-Jul-2000
 C:Accession: A35644; S13211
 R:Tashiro, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimonishi, M.; Shimizu, S.; Nakamura, Proc. Natl. Acad. Sci. U.S.A. 87, 3200-3204, 1990
 A:Title: Deduced primary structure of rat hepatocyte growth factor and expression of the A:Reference number: A35644; MUID:90222197
 A:Accession: A35644
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-728 <TAS>
 A:Cross-references: GB:P90102; GB:M32987; NID:q220766; PIDN:BA14133.1; PID:q220767
 A:Note: the authors translated the codon GAG for residue 70 as Gln, GAC for residue 417
 R:Okajima, A.; Miyazawa, K.; Kitamura, N.
 A:Title: Primary structure of rat hepatocyte growth factor and induction of its mRNA dur
 A:Reference number: S13211; MUID:91031482
 A:Accession: S13211
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-728 <OKA>
 A:Cross-references: EMBL:X54400; NID:956353; PIDN:CA38266.1; PID:94539554
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A:Note: does not have proteinase activity
 C:Superfamily: hepatocyte growth factor; kirtle homology; trypsin homology
 C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kirtle; py
 F:1-33/Domain: signal sequence #status predicted <SIG>
 F:56-495/Product: hepatocyte growth factor #status predicted <MAT>
 F:56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
 F:129-207/Domain: kirtle homology <KR1>
 F:212-289/Domain: kirtle homology <KR2>
 F:306-384/Domain: kirtle homology <KR3>
 F:382-470/Domain: kirtle homology <KR4>
 F:486-719/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
 F:496-719/Domain: trypsin homology <TRY>
 F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
 F:295,403,569,656/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:488-607/Disulfide bonds: #status predicted

Query Match 92.1%; Score 3802; DB 1; Length 728;
 Best local Similarity 90.4%; Pred. No. 1.4e-259;
 Matches 657; Conservative 41; Mismatches 27; Indels 2; Gaps 1;

1 MWTKLLPALLQHVLLHLLPRAIPYAGQRRRTTIEFKSAKTLTIKIDPALKIK 60
 2 MWGKTLPLVLLQHVLLHLLPRTIPYAGQRRRTTIEFKSAKTLTIKIDPALKIK 61
 61 TKRVNTADOCANCTRNKGIPTFCARFVDPKARQCLMPFNSMSGVKKFGEHFDLYE 120
 62 TKRVNSADECANCTRNKGIPTFCARFVDPKARQCLMPFNSMSGVKKFGEHFDLYE 121
 121 NKQYIRNCITIGKRSYGVSTFKSGIKCPWSSMTIPHEHSFLPSSYRGDLQENYCRNP 180
 122 NKQYIRNCITIGKRSYGVSTFKSGIKCPWSSMTIPHEHSFLPSSYRGDLQENYCRNP 181
 181 RGEDEGWCFTSNPEYRYEYCDIPQCEVECMCNCSYGLMDHTSSGICQWRMDQTP 240
 182 RGEDEGWCFTSNPEYRYEYCDIPQCEVECMCNCSYGLMDHTSSGICQWRMDQTP 241
 241 HRHKFLPERYPDKGPDNCRNDGPRPWCYTLDPTREYCAIKTCANTMNDTDPVL 300
 242 HRHKFLPERYPDKGPDNCRNDGPRPWCYTLDPTREYCAIKTCANTMNDTDPVL 301
 301 ETTECTGCGEGRGVNTIWMGIPCOQRWDQYRPHEDMTPEFKCKDLRENTCRNDGS 360
 302 ETTECTGCGEGRGVNTIWMGIPCOQRWDQYRPHEDMTPEFKCKDLRENTCRNDGS 361
 361 ESPWCTTDPNINIVGCSQIPNCDMSHGDCYGNCKNYGNLSQTRSGLTCSMDKME 420

362 ESPWCTTDPNINIVGCSQIPNCDMSHGDCYGNCKNYGNLSQTRSGLTCSMDKME 421
 421 DLHRIIFWEPDASKLNNYCRNPDADGAPMCYGNPLIPMDYCPISRCEDTPTIYNL 480
 422 DLHRIIFWEPDASKLNNYCRNPDADGAPMCYGNPLIPMDYCPISRCEDTPTIYNL 481
 481 DHPVTSCKATQOLRVNCLIPRTNIGMWSLRKRNKHTIGGSLIKESVLTAROCFPRN 539
 482 DHPVTSCKATQOLRVNCLIPRTNIGMWSLRKRNKHTIGGSLIKESVLTAROCFPRN 541
 540 -DLKDYEWMLGTHVHGSGDEKCOVNLVSQVYCPESGSDVLVKKLAPAYLDQFVSTID 598
 542 KDKMDYEWMLGTHVHGSGDEKCOVNLVSQVYCPESGSDVLVKKLAPAYLDQFVSTID 601
 601 LPSYCTIPEKTCISYGMWYTGILINADGGLRVNHLVIMGNKCSQHHQGVTLNESLTC 661
 659 LPNCGCTIPEKTCISYGMWYTGILINADGGLRVNHLVIMGNKCSQHHQGVTLNESLTC 658
 662 AGAERTGSGPCEGEGYGLVCEQHKRMVLYVPGRCALPNRPGIFVAVYAKWIK 718
 719 ILTVYKV 725
 722 VILTVKL 728

RESULT 4
 151283
 hepatocyte growth factor precursor - clawed frog
 N:Alternate names: hepatopietin A; scatter factor
 C:Species: Xenopus sp. (clawed frog)
 C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
 C:Accession: S15283
 R:Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiohawa, K.
 Mech. Dev. 49, 123-131, 1995
 A:Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in Xenopus
 A:Reference number: S15283; MUID:95267690
 A:Accession: S15283
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-710 <NAK>
 A:Cross-references: GB:S77422; NID:9998932; PIDN:AA34354.1; PID:9998933
 A:Note: the authors' translation for residue 458 (Thr) is inconsistent with the nucle
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A:Note: does not have proteinase activity
 C:Superfamily: hepatocyte growth factor; kirtle homology; trypsin homology
 C:Keywords: duplication; glycoprotein; growth factor; heterodimer; kirtle
 F:42-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
 F:115-193/Domain: kirtle homology <KR1>
 F:198-275/Domain: kirtle homology <KR2>
 F:289-367/Domain: kirtle homology <KR3>
 F:375-453/Domain: kirtle homology <KR4>
 F:478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
 F:478-709/Domain: trypsin homology <TRY>
 F:52,128,281,322,379,550,637,666/Binding site: carboxylate (Asn) (covalent) #status
 F:470-588/Disulfide bonds: #status predicted

Query Match 68.0%; Score 2806; DB 1; Length 710;
 Best local Similarity 66.1%; Pred. No. 1.5e-189;
 Matches 481; Conservative 107; Mismatches 120; Indels 20; Gaps 5;

1 MWTKLLPALLQHVLLHLLPRAIPYAGQRRRTTIEFKSAKTLTIKIDPALKIK 60
 1 MWTKLLPALLQHVLLHLLPRAIPYAGQRRRTTIEFKSAKTLTIKIDPALKIK 60
 61 TKRVNTADOCANCTRNKGIPTFCARFVDPKARQCLMPFNSMSGVKKFGEHFDLYE 120

```

Db 48 TRKNTTENCARKSCNNKGLPTCKAFADFNKIRCHMFSEFNTMSAGIKDKDYIDSEFLYE 107
QY 121 NKDYIRNCIIIGRSYKGTIVTITKSGIKCOPMSMTIPIHESFLIPSSYRGDLQENCRNP 180
Db 108 KQYIINDCJHGGSNRNGTRNTRKGLAQCPNMSMTIPIHESFLIPSTYRGDLKENCRNP 167
QY 181 RGEEGPWCFTSNPEVREYECDDIPQSEVCMTCNGESYGLMDHTESGKICQRMHDQRP 240
Db 168 KGESEGPWCFTSNPEVREYECDDIPQSEVCMTCNGESYGLMDHTESGKICQRMHDQRP 227
QY 241 HRHKLPERYPDKGFDDNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTNDTVDPL 300
Db 228 HKHKEFRPERYPDKGFDDNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTNDTVDPL 285
QY 301 ETEECTIOGEGYRGVNTIWMGICPQRMDSQYRPHEDMTPEHFCKKDLRENTCRNPDGS 360
Db 286 -TKDCKGSGEGYRGVNTIWMGICPQRMDSQYRPHEDMTPEHFCKKDLRENTCRNPDGS 344
QY 361 ESPWCFTTDPNIRIVGCSQIPNCDMSHGDCYRGNGKNYGNLSQTRSGLTCSMDPKNM 420
Db 345 ESPWCFTTDPNIRIVGCSQIPNCDMSHGDCYRGNGKNYGNLSQTRSGLTCSMDPKNM 404
QY 421 DLRHRTPEPDASKLNTNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTNDTVDPL 480
Db 405 DLRHRTPEPDASKLNTNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTNDTVDPL 464
QY 481 DHRVISCATKQDLRVYNGIPTRTNIGMWSLRVNRKHICGSLIKESWVLTAQCFPSR- 559
Db 465 DHRVISCATKQDLRVYNGIPTRTNIGMWSLRVNRKHICGSLIKESWVLTAQCFPSR- 523
QY 540 -DLKYEAMIGIHVGRGDECKQVLANYSOLVYEGPSDLVLMKARPAVLDDPFTSID 598
Db 524 IDLKYEAMIGIHVGRGDECKQVLANYSOLVYEGPSDLVLMKARPAVLDDPFTSID 562
QY 599 LNRGCTTPEKTSYSGYGGYGLNYDGLRYAHLYIMGNEKSCQHRHGKVTLNESEIC 658
Db 583 LNRGCTTPEKTSYSGYGGYGLNYDGLRYAHLYIMGNEKSCQHRHGKVTLNESEIC 642
QY 659 ACAERKIGSGPGGDDYGPVCEBOHKMRYLGVYVGRGALNPRRPIFRVAYAAWIHK 718
Db 643 ALGFRANIPCEBERDYGGLICEBENKTHLVQGYIIPGRGALQRPVIFRVAIYAKWIHK 702
QY 719 ILIYTKVP 726
Db 703 IMLTYKAP 710
RESULT 5
151285
hepatocyte growth factor/scatter factor - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: 151285
R:Steitl, A.; Stern, C.D.; Thery, C.; Ireland, G.W.; Aparicio, S.; Sharpe, M.J.; Cheraud
Development 121, 813-824, 1995
A:Title: A role for HGF/SF in neural induction and its expression in Hensen's node durin
A:Reference number: 151285; MUID:9537013
A:Accession: 151285
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-411 <STR>
A:Cross-references: GB:S77480; NID:9998675; PID:9998676
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
F:124-197/Domain: kringle homology <KRG>
F:202-279/Domain: kringle homology <KR2>
F:296-374/Domain: kringle homology <KR3>
Query Match 44.1%; Score 1821.5; DB 2: Length 411;
Best Local Similarity 73.1%; Pred. No. 14e-120;
Matches 309; Conservative 49; Mismatches 52; Indels 9; Gaps 2;
QY 1 MWVTKLLPALLQHVLLHLLPLAIPVNEGQRKRNRTIHEFKSAKTLIKIDPALKIK 60

```

```

Db 1 MMATQLLPALLH-----QLLLPITTPAAGKGRKRRRLHDYKKTGMLIKVNTLEVK 56
QY 61 TRKVNADQCANCRNTRKGLPTCKAFVFDKARKOCIMFPNMSGCVKKEFGEHDLYE 120
Db 57 TKLNTTEQCAKCRSRKGLSTFCKAFAFYDRAVRKCHWLSFNSLTNGVRKKQKHAFLE 116
QY 121 NKDYIRNCIIIGRSYKGTIVTITKSGIKCOPMSMTIPIHESFLIPSSYRGDLQENCRNP 180
Db 117 KQYIINDCJHGGSNRNGTRNTRKGLAQCPNMSMTIPIHESFLIPSTYRGDLKENCRNP 167
QY 181 RGEEGPWCFTSNPEVREYECDDIPQSEVCMTCNGESYGLMDHTESGKICQRMHDQRP 240
Db 172 RGEEGPWCFTTSPQARHEVCDDIPQSEVCMTCNGESYGLMDHTESGKICQRMHDQRP 221
QY 241 HRHKLPERYPDKGFDDNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTNDTVDPL 300
Db 232 HKHKEFRPERYPDKGFDDNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTNDTVDPL 285
QY 301 ETEECTIOGEGYRGVNTIWMGICPQRMDSQYRPHEDMTPEHFCKKDLRENTCRNPDGS 360
Db 292 ETTTCIOGEGYRGVNTIWMGICPQRMDSQYRPHEDMTPEHFCKKDLRENTCRNPDGS 351
QY 361 ESPWCFTTDPNIRIVGCSQIPNCDMSHGDCYRGNGKNYGNLSQTRSGLTCSMDPKNM 420
Db 352 ESPWCFTTDPNIRIVGCSQIPNCDMSHGDCYRGNGKNYGNLSQTRSGLTCSMDPKNM 410
RESULT 6
A47136
macrophage stimulating protein 1 precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 14-Nov-1997 #text_change 18-Jun-1999
C:Accession: A40331; B40331; A47136; A61395
R:Han, S.; Stuart, L.A.; Degen, S.J.F.
Biochemistry 30, 9768-9780, 1991
A:Title: Characterization of the DNFI552 locus on human chromosome 3: identification
A:Reference number: A40331; MUID:92002016
A:Accession: A40331
A:Molecule type: DNA
A:Residues: 1-711 <HA1>
A:Cross-references: GB:M74179
A:Accession: B40331
A:Molecule type: mRNA
A:Residues: 1-711 <HA2>
A:Cross-references: GB:M74178; NID:9183976; PIDN:AAA50165.1; PID:9183977
R:Yoshimura, T.; Yuhki, N.; Wang, M.H.; Skeel, A.; Leonard, E.J.
J. Biol. Chem. 268, 15461-15468, 1993
A:Title: Cloning, sequencing, and expression of human macrophage stimulating protein
A:Reference number: A47136; MUID:93340141
A:Accession: A47136
A:Molecule type: mRNA
A:Residues: 1-12, 'C', 14-622, 'F', 624-711 <YOS>
A:Cross-references: GB:111924; NID:9398037; PIDN:AAA59872.1; PID:9398038
A:Note: authors translated the codon TTT for residue 623 as Leu; parts of this sequen
R:Skeel, A.; Yoshimura, T.; Showalter, S.D.; Tanaka, S.; Appella, E.; Leonard, E.J.
J. Exp. Med. 177, 1227-1234, 1991
A:Title: Macrophage stimulating protein: purification, partial amino acid sequence, a
A:Reference number: A61395; MUID:91217635
A:Accession: A61395
A:Molecule type: protein
A:Residues: 230-247, 288-291, 'E', 293-295, 'X', 297-301, 'X', 303, 'E', 305, 'EX', 308-310, 326-
A:Experimental source: plasma
C:Genetics:
A:Gene: GDB:MS11; D3F15S2; DNFI5S2; HGFL
A:Cross-references: GDB:128833; OMIM:142408
A:Map position: 3p21-3p21.3
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-483, 484-711/Product: macrophage stimulating protein 1 #status predicted <MAT>
F:19-483/Domain: alpha chain #status predicted <ACH>

```


[illegible]

Db 354 TSREGLRNAFCQHOIPRGTIELVPECCYHGSGEOYRGVSQTKRKGVOQGH--SSEFTNPK 411

Qy 426 IFWEEDASK--LLENYCNRPDDDAHGWPCYTGANPLIFMDYCPISRCEDGTPPTIVNIDHP 483

Db 412 QFTPTSAFOAGILENFNCNRPDGDSDHGWCYTLDDPDLDFEYCALORCDDDDPPSILDPDQ 471

Qy 484 VI--SCA----KTQOLRVNCLIPRTNTGMMVSLRYR-NKIHGGSLIKESWVLTAROCF 536

Db 472 VFEEKGRKRVDSNKLRLVAGNH- GNSPWTVLSLRNROGHFGGSLVKEQWVLTARQCI 529

Qy 537 PS--RDLEKDYFAMLGIHVHNGRDEKCKQVLTNSOLVYPRPGSDLVLMKLARPAVLDFV 594

Db 530 WSCHEPLTGEVETWLTCTINQNPQPGANLQRRVPAKAGCAGSGLVLLKLEPPIILNHNH 589

Qy 595 STIDLPNVCCTIPKETSQSVYGMGYTGLINDGLRLVANHLYIMGENKCSQHNKRGVTLNE 654

Db 590 ALICPEPQYVVPPTCKEJAGWCSEISTSNNTVLHVASMNVISNOECNTKYRGIH--QE 647

Qy 655 SEICGAKKIGSGPEEGYGGPGLVCEQHKMMVLGVIYPRGGCALPNRPGIFGVAVYAK 714

Db 648 SEICQGLVPRVPGACEDYGGPGLACTYHDCWLVGLIIPNVCARPMRAIFTRVSVFVD 707

Qy 715 WIKKI 720

Db 708 WINKVM 713

RESULT 9

PLPG

plasmin (EC 3.4.21.7) precursor - pig (fragment)

N:Alternate names: plasminogen

N:Contains: miniplasminogen

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 07-Sep-1990 sequence revision 01-Nov-1996 #text_change 18-Jul-1997

C:Accession: S03733; S03737; A25834

R:Schaller, J.; Marti, T.; Roessellet, S.J.; Kaempfer, U.; Rickli, E.E.

R:Brinischolz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergler, W.; Manneberg, E.

Eur. J. Biochem. 114, 465-470, 1981

A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of human

A:Reference number: S03735; MUID:81212097

A:Accession: S03737

A:Molecule type: protein

A:Residues: 1-560 <SCH>

R:Brinischolz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergler, W.; Manneberg, E.

Eur. J. Biochem. 114, 465-470, 1981

A:Title: Determination of the complete amino-acid sequence of porcine miniplasminogen

A:Reference number: A25834; MUID:85203907

A:Accession: A25834

A:Molecule type: protein

A:Residues: 450-790 <MAR>

C:Function:

A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a

ns the walls of the graafian follicle; also activates the urokinase-type plasminogen

A:Pathway: fibrinolysis

C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolo

C:Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine pr

F:1-790/Product: plasminogen #status predicted <PRO>

F:1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLPR>

F:1-77/Domain: activation peptide #status predicted <APT>

F:78-560/Product: plasmin chain A #status predicted <ACH>

F:84-162/Domain: kringle homology <KR1>

F:166-243/Domain: kringle homology <KR2>

F:256-333/Domain: kringle homology <KR3>

F:358-435/Domain: kringle homology <KR4>

F:450-790/Product: miniplasminogen #status experimental <MIN>

F:461-540/Domain: kringle homology <KR5>

F:561-790/Product: plasmin chain B #status experimental <BCH>


```

158 YCRNPDNDENPCWCTTDPDKRYDYCDIPED-EDKCMHCSEGENEYKIAKTMGSRDCQAM 216
QY 236 DHQPHRHKFLPERYPKGFNDNCRNPDGPRRWCYTLDPHTREYCAITCADNTMND 295
Db 217 DQSPHAGYIPSPKFPNKLNMNCRNPDGPRRWCYTLDPHTREYCAITCADNTMND 270
QY 296 TDVPLE--TTTCIGOGEGYRGVNTIWNIGPCORWMDQYHEHDMPENEKCKDLRENY 353
Db 271 TTPPSSGPKYCKLGKGTGKNYGYAVYTESGHTCQRMSEQTTHKNNRTENPCKKLENY 330
QY 354 CRNPDGSESPWCTTDPNIRVGYCSQIPNCDMSH-----GQDCYRNG 396
Db 331 CRNPDGSESPWCTTDPNIRVGYCSQIPNCDMSH-----GQDCYRNG 389
QY 397 KNYGNLSQTSGLTCSMW-----DKN----- 418
Db 390 QSYRGTSSTYITGRKCSQSMSTPHRLKTPENPNAGLTNNYCRNPDADKSPWCYTDP 449
QY 419 ----- 418
Db 450 RVNREFCNLAKCSETPEDVPAPOAPGVENPPEADCMIGTKSTRGKATTVAGVPCQEW 509
QY 419 -MEDLHRIHEWEPDA---SKLNENYCRNPDADDAGPWCYTGNPILPMDYCIISRCGDTT 474
Db 510 AAGBPQHOSITPFTNPDQSGLENYCRNPDGDVNGPWCYTNNRKPFDYCDVPOCESS-- 567
QY 475 PTIYNLHPVISCATK-----QLRVNGIPIRTNT-GMAYSLRYRNKHGCGSLIKES 527
Db 568 -----FDCKRPKEPKCKSGRIYGGCVSKHSPWMOVSLRSHRHCQGLLISPK 617
QY 528 WVLTAAROCFSP-RDLKDEVAMLGIHDVHGRGDECKOVLNLSQLYVSEGSQDLVLMKLAR 586
Db 618 WVLTAAROCFSP-RDLKDEVAMLGIHDVHGRGDECKOVLNLSQLYVSEGSQDLVLMKLAR 675
QY 587 PAVLDDEVSTIDLPNYGCTIPEKTSQSVYGGWYGTGLINYDGLRVAHLIYINGNEKCSQNH 646
Db 676 PAITKEVYIPACLDPPNMAVARTECYITGMEGTGTGEGELKRAHLIYENKVCNNNE 735
QY 647 RGKTLNSELCAQAEKIGSGPCBEDYGGPLVCEQHKRMVGLVYVPRGCAIPRPGIF 706
Db 736 YLDRVAKTELCAGHLIGTDCSGDGGPLVCEKRYIIQGVYISWGLCARPKPGVY 795
QY 707 VRVAYYAKMIHK 718
Db 796 VRVSPYVPIWEE 807

RESULT 11
PLMS
plasmin (EC 3.4.21.7) precursor - mouse
N:Contains: angiotatin; plasminogen
C:Species: Mus musculus (house mouse)
C>Date: 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C:Accession: A38514; S48202; S48203
R:DeGen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
Genomics 8:49-61, 1990
A:Title: Characterization of the cDNA coding for mouse plasminogen and localization of
A:Reference number: A38514; MUID:91184812
A:Accession: A38514
A:Molecule type: mRNA
A:Residues: 1-812 <DEG>
A:Cross-references: GB:J04766; NID:g200402; PID:AAA50168.1; PID:g200403
R:Lijnen, H.R.; Van Hoet, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MUID:95010076
A:Accession: S48202
A:Molecule type: protein
A:Residues: 20-25 <LIU>
A:Accession: S48203
A:Molecule type: protein
A:Residues: 22-27 <LI2>
```

```

C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many
C:Comment: Plasminogen is converted into plasmin by plasminogen activators, both plas
C:Comment: immediately after dissociation from the clot. In the presence of the inhibitor, the act
C:Comment: the activation involves also removal of the activation peptide.
C:Comment: Strimvelis 1 (see PIR:KCMWS1) acts on plasminogen to produce angiotatin.
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C:Keywords: angiotensin inhibitor; blood; duplication; fibrinolysis; glycoprotein; h
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-812/Product: plasminogen #status predicted <PRO>
F:20-96/Domain: activation peptide #status predicted <APT>
F:79-466/Product: angiotatin #status predicted <AST>
F:97-581,582-812/Product: plasmin #status predicted <MAT>
F:97-581/Domain: chain A #status predicted <ACH>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:582-812/Domain: chain B #status predicted <BCB>
F:582-805/Domain: trypsin homology <TRY>
F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335
bonds: #status predicted
F:78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
F:136,308/Binding site: carboxylate (Asn) (covalent) #status predicted
F:466-467/Cleavage site: Thr-Val (plasminogen 1) #status predicted
F:581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F:624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 33.4%; Score 1379.5; DB 1; Length 812;
Best Local Similarity 34.4%; Pred. No. 3.6e-89;
Matches 288; Conservative 119; Mismatches 262; Indels 167; Gaps 22;

QY 15 VLLHLLPLIAPVYEGOKRRNTIHERKSAKTLIKIDPLAKIKTKVKNVA---DOCA 71
Db 7 ILLFLL-----KPGQ---GDSLDGYISTOGASIFSL-----TKQLAAGVSDCL 50.
QY 72 NCRTRNKGLPFTCKAFVDFPKARKOCLMPFPNSMSGVKPEFHEFDYENKRYINCIIG 131
Db 51 AKCEGE--TDVCRSFQYHSKQEOCVIAENSKTSITM--RDVYLEKRYVLSCKTG 106
QY 132 KGRSYKGVYSITKSGIKCPWSSMIPHEHSFLPSSYRGKLOENYCRNPDGEGSPWCT 191
Db 107 IONGRGYMSRKSGVACQKWAATPRHYVNYSPRHNNGLEENYCRNPDNDGQPMCT 166
QY 192 SNPEVREYVCDIPGSEVCEMTONGESYRGMDHTESGTCORWMDQPHRHKFLPERYP 251
Db 167 TDPDRKRYDCNIPCEEE-ECMYCSEKEYGKISKTMGIDCQAMPDSQSPHAGYIPAKFP 225
QY 255 DKGFNDNCRNPDGPRRWCYTLDPHTREYCAITCADNTMNDVPLETTEGIGOGGE 311
Db 226 SNLKNKNCCHNDGEPMPCTTDPTRKWEYCDIPRCT---TPPPSPYVQCLKGRGE 281
QY 312 GYRGTVNTIWNIGPCORWMDQYHEHDMPENEKCKDLRENYCRNPDGSESPWCTTDPN 371
Db 282 NRGTVSVYVSGKTCQRMSEQTTHKNNRTENPCKNLEENYCRNPDGTAATWCTTDSQ 341
QY 372 IRVGYCSQIPNCDMSHG-----ODCYRNGKNYGNLSQTSRGLTCSM 414
Db 342 LMEWYC-EIPSCSSASPQSDSVYVPEQTPVVOECYQSDQSGYSGTSITITKCKCS 400
QY 415 WDKNMDLHRIHFW---PDASKLNENYCRNPDADDAGPWCYTGNDLIPMDYCIISRCG 471
Db 401 WAMAFR--HRHSKTPENFPDAG-LENNYCRNPDG--KGMWCYTTPSVMEYCNLKR-S 455
QY 472 DTPPTIVNLHPVISCATK----- 486
```


Db	456	ETGGSVYEL--PYVSOERSGPDSDETDCMYGNGKDYRGKTAHTAAGTPOCGMAAQEPHRH	513
QY	487	-----CAKTKOL-----	493
Db	514	SIETPQTNPRADLEKNYCRNPBGDVNGPWCYTTPNPKLYDCIDPLCASASSFECKRPQV	573
QY	494	-----RVVNG-IPTPLTNIGMWSLYR--NKHCGSLIKESWVLTPAROCFSPRDLKD	543
Db	574	EPKPCGRVYVGGCVANHSHWPMOISLRTFRITQCHFCGGGLIAPENWLTALHCEKSSRPE	633
QY	544	-YEATLGIDHYHGRGDECKQOVLNVSQLVYPEGSDVLMLKARPAVLDDFVSTIDLPNY	602
Db	634	EKKVILGAHEHYEINGDL--VQEISVAKLILEPNNRNDIALKLSRPAITIDKYVAPACLPSP	691
QY	603	GCTIPEKTSQSVYMGVGTGLINVDGLRVANHLVIMGNEKSOHHNRGKVTLLNESEICGAE	662
Db	692	NYWVAADRITCIYIIMGWEGTQFFGAGRLKEADLPVLENNKCNKNVEYLYNNRYKSTELCAGOL	751
QY	663	KISGCPEDGYGCPFLVCEQHKKRMNVLYIVPBGCAINRPGIEFVAVAYAAKMIHK	718
Db	752	AGGVDSQQGGSGGGLVFEKDKYILLOGVTSMGLGCAIRPKPGVYVAVSRVRVDIER	807

RESULT 12

plasmin (EC 3.21.7) precursor [validated] - human
A:Alternate names: plasminogen precursor [misnomer]
N:Contains: angiotensin; microplasmin; plasminogen
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence_revision 02-Dec-1994 #text_change 15-Sep-2000
A:Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A04625; A04626
R:Peterson, T.E.; Marten, M.R.; Ichinose, A.; Davle, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibrinolytic system
A:Reference number: A35229; MUID:90202879
A:Molecule type: DNA
A:Accession: A35229
A:Residues: 1-810 <PEP>
A:Cross-references: GB:J05286; GB:M34276; NID:g190064; PIDN:AAA60113.1; PID:g387026
A:Experimental source: leukocyte; lung fibroblast
R:Margaret, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.; Tarantini, A.; Bressan, G.; Res. Commun. 173, 1013-1018, 1990
A:Title: Definition of the transcription initiation site of human plasminogen gene in 11p15.5
A:Reference number: I52242; MUID:91097523
A:Accession: I52242
A:Status: translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-16 <MAL1>
A:Cross-references: GB:M62890; NID:g190092; PIDN:AAA36454.1; PID:g553613
R:Jorsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
FEBS Lett. 213, 254-260, 1987
A:Title: Molecular cloning and characterization of a full-length cDNA clone for human plasminogen activator
A:Reference number: A26646; MUID:87162490
A:Accession: A26646
A:Molecule type: mRNA
A:Residues: 1-471, 'D', 473-810 <FOR>
A:Cross-references: GB:X05199; NID:g35530; PIDN:CA28831.1; PID:g35531
A:Experimental source: liver
R:Malinowski, D.P.; Sadler, J.E.; Davle, E.W.
Biochemistry 23, 4243-4250, 1984
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and rat plasminogen activator
A:Reference number: I45961; MUID:85023311
A:Accession: I62738
A:Status: translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 292-471, 'D', 473-810 <MAL2>
A:Cross-references: GB:K02922; NID:g190112; PIDN:AAA60124.1; PID:g387031
A:Accession: I84609
A:Status: translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 367-419 <MAL3>
A:Cross-references: GB:K02921; NID:g190110; PIDN:AAA60123.1; PID:g190111
R:Brundish, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergler, W.; Manneberg, M.;

Eur. J. Biochem. 114, 465-470, 1981

A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human plasminogen and human fibrinogen

A:Reference number: S03735; MUID:81212097

A:Accession: S03735

A:Molecule type: protein

A:Residues: 20-71, 'E', 73-76 <BRU>

R:Sortrup-Jensen, L.; Petersen, T.E.; Magnusson, S. submitted to the Atlas, July 1977

A:Reference number: A00929

A:Accession: A00929

A:Molecule type: protein

A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>

R:Wiman, B.

Eur. J. Biochem. 76, 129-137, 1977

A:Title: Primary structure of the B-chain of human plasmin.

A:Reference number: A04627; MUID:77225245

A:Accession: A04627

A:Molecule type: protein

A:Residues: 581-810 <WII>

R:Wiman, B.; Wallen, P.

Eur. J. Biochem. 50, 489-494, 1975

A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human plasminogen

A:Reference number: A04625; MUID:75093329

A:Accession: A04625

A:Molecule type: protein

A:Residues: 20-50, 'Q', 51-71, 'E', 73-85, 87-100 <WII>

R:Wiman, B.; Wallen, P.

Eur. J. Biochem. 58, 539-547, 1975

A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen

A:Reference number: A04626; MUID:76043692

A:Accession: A04626

A:Molecule type: protein

A:Residues: 483-507, 'E', 509-604 <WII>

R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.

J. Biol. Chem. 248, 1631-1633, 1973

A:Title: The primary structure of human plasminogen. II. The histidine loop of human plasminogen

A:Reference number: A92125; MUID:73149248

A:Contents: annotation; active site

R:Groskopf, W.R.; Summaria, L.; Robbins, K.C.

J. Biol. Chem. 244, 3590-3597, 1969

A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of human plasminogen

A:Reference number: A92048; MUID:69234739

A:Contents: annotation; active site

R:Trexlter, M.; Vall, Z.; Patchy, L.

J. Biol. Chem. 257, 7401-7406, 1982

A:Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen

A:Reference number: A92382; MUID:82213905

A:Contents: annotation; omega-aminocarboxylic acid binding sites

R:Vall, Z.; Patchy, L.

J. Biol. Chem. 259, 13690-13694, 1984

A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential for fibrin binding

A:Reference number: A92458; MUID:85054794

A:Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site

R:Gao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.

J. Biol. Chem. 271, 29461-29467, 1996

A:Title: Kringle domains of human angiotensin. Characterization of the anti-proliferative activity of the kringle domain

A:Reference number: A58811; MUID:97062211

A:Contents: annotation

R:Liinen, H.R.; Uggwu, F.; Bini, A.; Collen, D.

Biochemistry 37, 4699-4702, 1998

A:Title: Generation of an angiotensin-like fragment from plasminogen by stromelysin-1

A:Reference number: A58812; MUID:9548733

A:Contents: annotation

R:Tulinsky, A.; Mulichak, A.M.

submitted to the Brookhaven Protein Data Bank, July 1991

A:Reference number: A51341; PDB:1PK4

A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454

R:Tulinsky, A.; Wu, T.P.

submitted to the Brookhaven Protein Data Bank, July 1991

A:Reference number: A51488; PDB:2PK4

A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454

R:Wu, T.P.; Tulinsky, A.

submitted to the Brookhaven Protein Data Bank, August 1993
 A:Reference number: A51911; PDB:IPKR
 A:Contents: annotation: X-ray crystallography, 2.48 angstroms, residues 102-181
 R:Padmanabhan, K.; Tulinsky, A.
 Submitted to the Brookhaven Protein Data Bank, April 1994
 A:Reference number: A52408; PDB:IPMK
 A:Contents: annotation: X-ray crystallography, 2.25 angstroms, residues 377-454
 R:Tulinsky, A.; Mathews, I.I.
 Submitted to the Brookhaven Protein Data Bank, December 1995
 A:Reference number: A65244; PDB:ICEA
 A:Contents: annotation: X-ray crystallography, 2.1 angstroms, residues 102-181
 R:Tulinsky, A.; Mathews, I.I.
 Submitted to the Brookhaven Protein Data Bank, December 1995
 A:Reference number: A65245; PDB:ICEB
 A:Contents: annotation: X-ray crystallography, 2.1 angstroms, residues 102-181
 R:Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
 Biochemistry 30, 10576-10588, 1991
 A:Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å
 A:Contents: annotation
 A:Reference number: A58819; PMID:92031502
 R:Mu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
 Biochemistry 30, 10589-10594, 1991
 A:Title: The refined structure of the epsilon-aminocaproic acid complex of human plasmin
 A:Contents: annotation
 A:Reference number: A58818; PMID:92031503
 R:de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M.
 Biochemistry 31, 270-279, 1992
 A:Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2 Å
 A:Reference number: A39483; PMID:92118803
 A:Contents: annotation: X-ray crystallography, 2.4 angstroms
 R:Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.
 Submitted to the Brookhaven Protein Data Bank, June 1995
 A:Reference number: A65980; PDB:IKRN
 A:Contents: annotation: X-ray crystallography, 1.67 angstroms, residues 376-454
 R:Rejante, M.; Llinas, M.
 Submitted to the Brookhaven Protein Data Bank, August 1996
 A:Reference number: A65803; PDB:IHPJ
 A:Contents: annotation: conformation by (1)H-NMR, residues 103-181
 R:Rejante, M.; Llinas, M.
 Submitted to the Brookhaven Protein Data Bank, August 1996
 A:Reference number: A65804; PDB:IHPK
 A:Contents: annotation: conformation by (1)H-NMR, residues 103-181
 R:Rejante, M.R.; Llinas, M.
 Eur. J. Biochem. 221, 927-937, 1994
 A:Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
 A:Reference number: A43645; PMID:94237157
 A:Contents: annotation: conformation by (1)H-NMR, residues 96-184
 R:Rejante, M.R.; Llinas, M.
 Eur. J. Biochem. 221, 939-949, 1994
 A:Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen
 A:Reference number: A58817; PMID:94237158
 A:Contents: annotation: conformation by (1)H-NMR
 C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many of
 C:Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKHU and
 C:Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHA2) immediately after
 C:Comment: Microplasmin is formed by autocatalytic cleavage of plasmin under artificial conditions
 C:Comment: Streptolysin 1 (see PIR:KCHU1) acts on plasminogen to produce angiotensin. The
 C:Genetics:
 A:Gene: GDB:PLG
 A:Cross-references: GDB:119498; OMIM:173350
 A:Map position: 6q26-6q27
 A:Introns: 11/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 529/2
 C:Function:
 A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety
 A:Pathway: fibrinolysis
 C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
 C:Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydrolysis
 F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
 F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-810/Product: plasminogen #status experimental <PRO>
 F:20-96/Domain: activation peptide #status experimental <APT>
 F:79-46/Product: angiotensin #status experimental <AST>
 F:97-580,581-810/Product: plasmin #status experimental <MAT>
 F:97-580/Domain: plasmin chain A #status experimental <CHA>
 F:103-181/Domain: kringle homology <KR1>
 F:185-262/Domain: kringle homology <KR2>
 F:275-352/Domain: kringle homology <KR3>
 F:377-454/Domain: kringle homology <KR4>
 F:481-560/Domain: kringle homology <KR5>
 F:530-580,581-810/Product: microplasmin #status experimental <MPT>
 Query Match 33.2%; Score 1371; DB 1; Length 810;
 Best Local Similarity 34.4%; Pred. No. 1,4e-88;
 Matches 292; Conservative 118; Mismatches 257; Indels 182; Gaps 22;
 QY 12 LQVHLHLPLPAIPYAEQGRKRRNTIHERKSAKTLTKIDPAIKTKK---VNTAD 68
 DB 1 MEHEVLLLL-LFLKSGGE-----PLDDYVNTQGSLSFSV-----TKQLGAGSIE 47
 QY 69 QCANRCRRNGLPRTCKAFAPFDKARKOCIMPRFMSNGVKKERGFEDLEYENKDYLRNC 128
 DB 48 ECAAKCEDE--EFTCAFYHNSKEQOCVMAENRKSIIIR--KQVLEKKVYLSEC 103
 QY 129 IIGKRSYKGVSTITKSGIKQCPWSSMIPIHESPLPSSRGKDLQENYCRNPRGEGGPW 188
 DB 104 KTGNGKRYRGMSKTNGITCQKWSSTSPHRRSPATNHPSEGLENYCRNPDNDPGPW 163
 QY 189 CFTSNPEVRREVCIDPQCEVEYECMTGNGESTYRGLMDHTSGKTCQRMDDHPTPHRKLP 248
 DB 164 CYTDPPEKRYDCIDILCEE-ECMHCSGENDGKISTMSGLQCAWDSPHAGYIPS 222
 QY 249 RYPRKGFDDNCRNPDGCRPWCYTLDPRHPRMECAIKTCQADNTMDPTDPLE--TTECI 306
 DB 223 KFPKKNLKKNTCRNPDDELAPMCFTTDPNKRKWLCDIPRCT-----TPPSSCPITQCL 276
 QY 307 QGQEGYRGTYNTIWNIGIPCRQMDQYRPHEDMTPEHFKCKDLRENYCRNPDSESPCF 366
 DB 277 KGTGENRGNAVAVVSGHTQGMASQTPRHNRPRPNPCNDENYCRNPDGKRAVWCH 336
 QY 367 TTDPNIVGCSQIPLNCDSH-----GDCYRGNKRNKTNISQTRSG 409
 DB 337 TTNSQVMEYC-KIPSCDSSPVSTEO LAPAPRLTAVVQDCYHGDSQYRGTSSTTTTG 395
 QY 410 LTGSMV-----DK----- 417
 DB 396 KKCQSWSSMTPRHQKRPENTPNAGLTMNCRNPDADKGPWCFTTDPVMEYCNLKKCS 455
 QY 418 -----NME-----DLARRHF 427
 DB 456 GTEASVAPRPVLLPVRVETPSEEDCMFNGCKYRGKRAITVTGTPCODMAAOEPRHHSI 515
 QY 428 WEPDASK---LNNENYCRNPDDAHGPCYTGNDPLIPWDYCPISRCESGDTPTIYNDHPV 484
 DB 516 FTPTNPBAGLEKKNYCRNPDQVGGPCYTTNPKRIDYCDVCAAA-----PS 564
 QY 485 ISCAKTK-----QLRVNGLPTPTNTI-CMNVSLRYR-NKHIGGSLIKESWVLTARQCF 536
 DB 565 FDGCKPVEPRKPCRGVAVGGCVVAHPHSPMPOVSLTRFGMIFCGGTLISPEWVLTAAHCL 624
 QY 537 -----PSRDLKDYEWAMLGIDHVGGRDEKCKQVLYNSQIYVGPESGLVLMKLARPAVL 591
 DB 625 EKSPRPS-----SYKVLGAHDEVNL--EPHQEILEVRLFLPFRKRIKDLSSPAVIT 678
 QY 592 DEVSFTLDPNYGCTIPERTKTSYVGMGYTGLINYDGLRVAHLIYMGNEKCSQHHRRKVT 651
 DB 679 DKVLPALPSPVYVAVDRTECFITGMGETQGTFGAGLKEQLDVIENKVCNRRFEYNGR 738
 QY 652 LNSEIICAGAEKIGSGCEDGYPGLVCEQKMMVGLYVPGGCAIPNPGIFGVVAY 711
 DB 739 VQSEILCAGHLAGTSDSCQDGSGLPVCFEKDYITLQGVTSWGLGCAKPNKPGVYVRSR 798
 QY 712 YAKWIKRII 720

Db 799 FVTWIEGV 807

RESULT 13

B30848
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 22-Jun-1999
C:Accession: B32869; B30848
R:Tominson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A:Reference number: A32869; MUID:89174660
A:Accession: B32869
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-810 <TOM>
A:Cross-references: GB:J04697; NID:g342272; PIDN:AAA36901.1; PID:g342273
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-9/Domain: signal sequence #status predicted <SIG>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:581-803/Domain: trypsin homology <TRY>
F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,324-342,665,760/Active site: His, Asp, Ser #status predicted

Query Match 33.2%; Score 1369; DB 2; Length 810;

Best Local Similarity 33.8%; Pred. No. 2e-88;

Matches 286; Conservative 127; Mismatches 257; Indels 176; Gaps 21;

Db 12 LQVHLHLLELPVIAPEAGORRRNTIHEFKSAKTLIKIDPALKIKRK--VMTAD 68
1 MEHKEVYLLLL-LFKSGGE-----PLDLYVNTKGASLSI-----IKKGAGASIE 47
QY 69 QCANRCRNRKGLPTCAFYFDKARKKCLMFPPNSMSSGVKKEFGHEFDLYENKDYTRNC 128
48 ECAKACEEEE--ETCSFQYHSKEDQCVIAENRKSIVFRM--RDVYLFKKVYVLESEC 103
QY 129 IICKSGSYKTSVITKSGICOPWMSWIRPHEHSFLPSSYSGKDIOENYCNPRDEEGPW 188
104 KTGNGKNYRGTMTKRTKRTGTCQAKSSISPHRPTFSPTAHSEGLEENYCRNPDNDGGGPW 163
189 CFTSNPEVRYEVDIPQCSVEECMTGSEYRGLMDHTESGICQRMWDQTPHNRKFLPE 248
164 CYTTDGEEREDYCDIPEC--EDECMHCSGSENYDKISKTMSGLECOAMDQSPAHAGYIPS 222
QY 249 RYDDKGGDDNYCNPDPGQRPWCYTLDPHRMVECAIKTADNTMNDTVPLE--TTETCI 306
223 KFNKMLKKNYCNPNPGEERPMCFETDPNKRMLCDIPRT-----TPPSSGPTYOQL 276
QY 307 QGQEGEYRGTVNTWNGIIPQCRWDQYRPHEDMTPENFKCKDLRENYCRPDGSESPWCF 366
277 KGTGENYRGDAVAVTSGHCHGWSAQTPIHNTPEPFCKNIDENYCRNPDEKAPWCY 336
QY 367 TTDPIRNVGICSGIIPNCDSH-----GDCYRGNKNYMGNLISQTRSG 409
337 TNSOVRMEYC-KIPSCSPVSTPELDPTAPPELVVQECYHGGQSYRGSSITTTTG 395
QY 410 LFCSMYDKMMDLHR-----HIFWE----- 429
Db 396 KKCQSSMSTPIHHEKTPENFPNAGLTMYNCRNPADKGRWCFTTDPVSRWEYCNLKCS 455
QY 430 -----PDA----- 432
Db 456 GTGSAVAPPAVPAQADATPSEEDCMFNGKGYRGKAKATYVTGTCQEMAQDEPHSHRI 515

QY 433 -----SKLNENYCRNPDDDAGHPWCYTGNPILLIPWDCYFISRCGDDTPTIVNLDPV 484
Db 516 FTPEINPRAGLEKNYCRNPDDGVDGPPWCYTTPNPKLFYCDVCPQCAASS----- 564
QY 485 ISCAKTK-----QLRYVNG--IPIRTNIGWWSLRYR-NKHCGSLIKESWLTAROCF 536
Db 555 FDCCGPOVEPKKCPGRVAVGCVAPHSMPWQISLTRILGMHFCGGTLLSPWVLAACL 624
QY 537 PSRLDKL-YEAMLGTH-DVHGKDEKCKQVNLVNSQVYGPBGSDVLMKLARPVLDDEV 594
Db 625 EKSSRPSFYKYLGAHREHVH---LEPHQDELVEKSMFSEPARADIALKLSSPAITDKV 681
QY 595 STIDLPNKGCIPEKTSQSVYGMWYTGILNYDGLIRVAHLTYMGNEKCSOHNRKVTLINE 654
Db 682 IPACLPSPNYVADRTCEFTIGMETQGTGAGLILKEARLPIENKVCNRYEFLNGYVKT 741

QY 655 SEICGAEKIGSGPOEGYGGPLVCEQHKRMVGLVYPRGCAIPNRGTFVAVAYYAK 714
Db 742 TELCAGHLAGTDSGQDSSGSLVCFEKDYILLOGVTSWGLGCAKAPMKPGVYVAVSRFVT 801
QY 715 WIKRII 720
Db 802 WIEGV 807

RESULT 14

I46260
C:Species: Erinaceus europaeus (western European hedgehog)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C:Accession: I46260
R:Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; F
J. Biol. Chem. 270, 24004-24009, 1995
A:Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipop
A:Reference number: I46259; MUID:96025778
A:Accession: I46260
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-810 <LAW>
A:Cross-references: EMBL:U03171; NID:g1046360; PID:g1046361
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C:Keywords: hydrolase; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:379-456/Domain: kringle homology <KR4>
F:482-561/Domain: kringle homology <KR5>
F:582-803/Domain: trypsin homology <TRY>

Query Match 32.8%; Score 1353.5; DB 2; Length 810;

Best Local Similarity 35.1%; Pred. No. 2.4e-87;

Matches 277; Conservative 110; Mismatches 255; Indels 147; Gaps 16;

QY 61 TKR-----VNTADCANCTNRKGLPTCAFYFDKARKKCLMFPPNSMSSGVKKEFGHEHD 117
Db 37 TKQOLSVGSEECAYVC--EKETSFICRSFOYHSKEDQCVIAENKSTPVLRM--RDVI 92
QY 118 LYENKDYIRNCITIGKRSYKTSVITKSGICQPMSSMIRPHEHSFLPSSYRGKDIOENY 177
Db 93 LFEKMYLYSECKVGNKGYRGTYSKRTGLTQKMAELPHKRFEPDENPSEGLDQNYC 152
QY 178 RNRGEGGPPWCTSNPEVRYEVDIPQCSVEECMTGSEYRGLMDHTESGICQRMWDH 237
Db 153 RNPNDPKGPPWCYTMDPEVRYEYCELIQC--EDECMHCSGQNYGVKISRMSGLECOPMDS 211
QY 238 QPFRHKKFLPERYDQGFDPDNCRNPDGQRPWCYTLDPHRMVECAIKTADNTMNDTD 297
Db 212 QTPRHGFTPSKFPKSNLKNYCNPNPGEERPMCFETDPNKRMLCDIPRT-----TP 265
QY 298 VPLE--TTETCIOGQEGYRGTVNTWNGIIPQCRWDQYRPHEDMTPENFKCKDLRENYCR 355

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2002, 17:46:19 ; Search time 13.55 Seconds
(without alignments)
2080.283 Million cell updates/sec

Title: US-09-423-516-2
Perfect score: 4126
Sequence: 1 MWVTKLLPALLQHVLLHL.....VAYYAKWIKITITLYKVPQS 728

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Fetched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4126	100.0	728	1 HGF_HUMAN	P14210 homo sapien
2	3804	92.2	728	1 HGF_MOUSE	Q08048 mus musculu
3	3802	92.1	728	1 HGF_RAT	P17945 rattus norv
4	1652	40.0	711	1 HGFL_HUMAN	P26927 homo sapien
5	1602	38.8	716	1 HGFL_MOUSE	P26928 mus musculu
6	1446.5	35.1	790	1 PLMN_PIG	P06867 sus scrofa
7	1408.5	34.1	812	1 PLMN_BOVIN	P06868 bos taurus
8	1379.5	33.4	810	1 PLMN_MOUSE	P20918 mus musculu
9	1370	33.2	810	1 PLMN_HUMAN	P00747 homo sapien
10	1369	33.2	810	1 PLMN_MACMU	P12545 macaca mula
11	1353.5	32.8	810	1 PLMN_ERIEU	Q29485 erinaceus e
12	1212.5	29.4	4548	1 APOA_HUMAN	P08319 homo sapien
13	1188.5	28.8	1420	1 APOA_MACMU	P14417 macaca mula
14	583	14.1	343	1 PLMN_SHEEP	P81486 ovis aries
15	559.5	13.6	559	1 TPA_RAT	P19637 rattus norv
16	552.5	13.4	338	1 PLMN_HORSE	P80010 equus cabal
17	551.5	13.4	559	1 TPA_MOUSE	P11214 mus musculu
18	548.5	13.3	333	1 PLMN_CANFA	P80009 canis fami
19	543	13.2	625	1 THRB_BOVIN	P00735 bos taurus
20	539.5	13.1	566	1 TPA_BOVIN	Q28198 bos taurus
21	532.5	12.9	562	1 TPA_HUMAN	P00750 homo sapien
22	531.5	12.9	618	1 THRB_MOUSE	P19321 mus musculu
23	527.5	12.8	622	1 THRB_HUMAN	P00734 homo sapien
24	518	12.6	617	1 THRB_RAT	P18292 rattus norv
25	510.5	12.4	603	1 FA12_CAVPO	Q04962 cavia porce
26	497.5	12.1	653	1 HGFA_MOUSE	Q04962 mus musculu
27	480.5	11.6	625	1 HGFA_HUMAN	Q04756 homo sapien
28	449	10.9	325	1 PLMN_PETMA	P33574 petromyzon
29	449	10.9	615	1 FA12_HUMAN	P00748 homo sapien
30	426.5	10.3	593	1 FA12_BOVIN	P98140 bos taurus
31	393	9.5	433	1 UROK_BOVIN	Q05589 bos taurus
32	350	9.5	431	1 UROK_HUMAN	P00749 homo sapien
33	388	9.4	433	1 UROK_MOUSE	P06869 mus musculu

34	386	9.4	442	1 UROK_PIG	P04185 sus scrofa
35	385	9.3	434	1 UROK_CHICK	P15120 gallus gall
36	382.5	9.3	432	1 UROK_RAT	P29598 rattus norv
37	382.5	9.3	433	1 UROK_PAPCY	P16227 papio cynoc
38	375	9.1	1035	1 ENTK_BOVIN	P98072 bos taurus
39	373.5	9.1	477	1 URT2_DESRO	P15638 desmodus ro
40	373	9.0	477	1 URT1_DESRO	P98119 desmodus ro
41	372	9.0	273	1 MCT7_MOUSE	Q02844 mus musculu
42	369	8.9	431	1 URT6_DESRO	P98121 desmodus ro
43	358	8.7	273	1 MCT7_RAT	P27435 rattus norv
44	358	8.7	394	1 URT6_DESRO	P27435 desmodus ro
45	357	8.7	1069	1 ENTK_MOUSE	P97435 mus musculu

ALIGNMENTS

RESULT 1
HGF_HUMAN STANDARD: PRT; 728 AA.
ID HGF_HUMAN
AC P14210; O99BU6; O99BU19;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
DE (Hepatopoietin A).
CN HGF OR HPTA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91340155; PubMed=1831432;
RA Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;
RT "Organization of the human hepatocyte growth factor-encoding gene.";
RT Gene 102:213-219(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=89392017; PubMed=2528952;
RA Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,
RA Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,
RT Gonda E., Daikuhara Y., Kitamura N.;
RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte
RT growth factor.";
RT Biochem. Biophys. Res. Commun. 163:967-973(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=leukocyte;
RX MEDLINE=91025062; PubMed=2145836;
RA Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T.,
RA Asami O., Hagiya M., Nakamura T., Shimizu S.;
RT "Isolation and expression of cDNA for different forms of hepatocyte
RT growth factor from human leukocyte.";
RT Biochem. Biophys. Res. Commun. 172:321-327(1990).
RN [4]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.
RC TISSUE=liver;
RX MEDLINE=9006676; PubMed=2531289;
RA Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,
RA Sugimura A., Tashiro K., Shimizu S.;
RT "Molecular cloning and expression of human hepatocyte growth factor.";
RT Nature 342:440-443(1989).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=embryonic fibroblast;
RX MEDLINE=91334393; PubMed=1831266;
RA Weidner K.M., Arakaki N., Hartmann G., Vandekerckhove J., Weingart S.,
RA Riederer H., Fomatsch C., Tsubouchi H., Hishida T., Daikuhara Y.,
RA Birchemer W.;
RT "Evidence for the identity of human scatter factor and human
RT hepatocyte growth factor.";

Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
 [6] SEQUENCE FROM N.A.
 RA Courtney L., Elliot G., Angell S.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RX SEQUENCE OF 249-695 FROM N.A.
 RA MEDLINE=91369928; PubMed=1832556;
 RA Miyazawa K., Kitamura A., Kitamura N.;
 RT "Structural organization and the transcription initiation site of the
 human hepatocyte growth factor gene.";
 RL Biochemistry 30:9170-9176(1991).
 RN [8]
 RP SIGNAL SEQUENCE CLEAVAGE SITE.
 RX MEDLINE=91207365; PubMed=1826837;
 RA Yoshiyama Y., Arakaki N., Naka D., Takahashi K., Hirano S., Kondo J.,
 RA Nakayama H., Gohda E., Kitamura N., Tsubouchi H., Ishii T.,
 RA Hishida T., Dakuwara Y.;
 RT "Identification of the N-terminal residue of the heavy chain of both
 native and recombinant human hepatocyte growth factor.";
 RL Biochem. Biophys. Res. Commun. 175:660-667(1991).
 RN [9]
 RP CARBOHYDRATE-LINKAGE SITE 476.
 RX MEDLINE=9139192; PubMed=1482348;
 RA Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,
 RA Nakamura T., Shimizu S.;
 RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
 on the alpha chain.";
 RL Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
 RN [10]
 RP MUTAGENESIS.
 RX MEDLINE=92331602; PubMed=1321034;
 RA Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
 RA Baker J.B., Godowski P.J.;
 RT "Structure-function analysis of hepatocyte growth factor:
 identification of variants that lack mitogenic activity yet retain
 high affinity receptor binding.";
 RL EMBO J. 11:2503-2510(1992).
 RN [11]
 RP STRUCTURE BY NMR OF 31-127.
 RX MEDLINE=98154323; PubMed=9493272;
 RA Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,
 RA Rubin J.S., Bottaro D.P., Byrd R.A.;
 RT "The solution structure of the N-terminal domain of hepatocyte growth
 factor reveals a potential heparin-binding site.";
 RL Structure 6:109-116(1998).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.
 RX MEDLINE=99036858; PubMed=9817840;
 RA Ullrich M., Lokker N.A., Godowski P.J., de Vos A.M.;
 RT "Crystal structure of the NK1 fragment of human hepatocyte growth
 factor at 2.0-A resolution.";
 RL Structure 6:1383-1393(1998).
 CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
 HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS
 AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
 CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
 DISULFIDE BOND.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
 CC =====
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC =====
 CC EMBL: D90334; BAA14348.1; -
 CC EMBL: D90318; BAA14348.1; JOINED.
 DR

DR EMBL: D90319; BAA14348.1; JOINED.
 DR EMBL: D90320; BAA14348.1; JOINED.
 DR EMBL: D90322; BAA14348.1; JOINED.
 DR EMBL: D90323; BAA14348.1; JOINED.
 DR EMBL: D90324; BAA14348.1; JOINED.
 DR EMBL: D90325; BAA14348.1; JOINED.
 DR EMBL: D90326; BAA14348.1; JOINED.
 DR EMBL: D90327; BAA14348.1; JOINED.
 DR EMBL: D90328; BAA14348.1; JOINED.
 DR EMBL: D90329; BAA14348.1; JOINED.
 DR EMBL: D90330; BAA14348.1; JOINED.
 DR EMBL: D90331; BAA14348.1; JOINED.
 DR EMBL: D90332; BAA14348.1; JOINED.
 DR EMBL: D90333; BAA14348.1; JOINED.
 DR EMBL: M29145; AAA52650.1; -
 DR EMBL: M60718; AAA52648.1; -
 DR EMBL: X16323; CAA34387.1; -
 DR EMBL: M73239; AAA64239.1; -
 DR EMBL: M73240; AAA64297.1; -
 DR EMBL: AC004960; AAC71655.1; -
 DR EMBL: M75983; AAG53460.1; -
 DR EMBL: M75972; AAG53460.1; JOINED.
 DR EMBL: M75973; AAG53460.1; JOINED.
 DR EMBL: M75974; AAG53460.1; JOINED.
 DR EMBL: M75975; AAG53460.1; JOINED.
 DR EMBL: M75976; AAG53460.1; JOINED.
 DR EMBL: M75977; AAG53460.1; JOINED.
 DR EMBL: M75978; AAG53460.1; JOINED.
 DR EMBL: M75979; AAG53460.1; JOINED.
 DR EMBL: M75980; AAG53460.1; JOINED.
 DR EMBL: M75981; AAG53460.1; JOINED.
 DR EMBL: M75982; AAG53460.1; JOINED.
 DR PIR: JH0579; JH0579.
 DR PIR: S06794; S06794.
 DR PDB: 2HGF; 24-JUN-98.
 DR PDB: 1BHT; 18-NOV-98.
 DR MEROPS; S01.976; -
 DR GlycoSuiteDB; P14210; -
 DR MIM; 142409; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00051; kringle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; TRYF_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS0070; KRINGLE_2; 4.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR Growth factor; Kringle; Glycoprotein; Serine protease homolog;
 KW Signal; 3d-structure.
 FT SIGNAL 1 31
 FT CHAIN 32 494
 FT CHAIN 495 728
 FT MOD_RES 32 32
 FT DOMAIN 32 127
 FT DOMAIN 128 206
 FT DOMAIN 211 288
 FT DOMAIN 305 383
 FT DOMAIN 391 469
 FT DOMAIN 495 728
 FT DISULFID 70 96
 FT DISULFID 74 84
 FT DISULFID 128 206
 FT DISULFID 149 189
 FT DISULFID 177 201
 FT DISULFID 487 604
 INTERCHAIN (BY SIMILARITY).
 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
 HEPATOCYTE GROWTH FACTOR BETA CHAIN.
 PYROLIDONE CARBOXYLIC ACID.
 PAF.
 KRINGLE 1.
 KRINGLE 2.
 KRINGLE 3.
 KRINGLE 4.
 SERINE PROTEASE-LIKE.

Query Match 100.0%; Score 4126; DB 1; Length 728;
 Best Local Similarity 100.0%; Pred. No. 6.5e-301;
 Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWVTKLLPALLQVHLLHLLLPALPAVAGGKRRNRTHEFFKSAKTTLIKIDPALKIK 60
 DB 1 MWVTKLLPALLQVHLLHLLLPALPAVAGGKRRNRTHEFFKSAKTTLIKIDPALKIK 60
 QY 61 TKRVNADOCANCTRNKGLPFCKAFVDDKARKOCLMPFNSSGVKKEFHEPDLYE 120
 DB 61 TKRVNADOCANCTRNKGLPFCKAFVDDKARKOCLMPFNSSGVKKEFHEPDLYE 120
 QY 121 NKDYINNCIIIGKRSYKGYSTIKSGIKCOPWSSMI PHEHSFPPSSYRGKDLQENYCRNP 180
 DB 121 NKDYINNCIIIGKRSYKGYSTIKSGIKCOPWSSMI PHEHSFPPSSYRGKDLQENYCRNP 180
 QY 181 RGEESGPMCFSTSNPEVREYECVDPQCSSEVEMTCNGESYRGMLDHTESGKICQRMHDQTP 240
 DB 181 RGEESGPMCFSTSNPEVREYECVDPQCSSEVEMTCNGESYRGMLDHTESGKICQRMHDQTP 240
 QY 241 HRKELPERYPRDGPNDNCRNPDGP RPWCYTLDPHTREYCAITCANTMNDVPL 300
 DB 241 HRKELPERYPRDGPNDNCRNPDGP RPWCYTLDPHTREYCAITCANTMNDVPL 300
 QY 301 ETTECIQGGEGYRGTVNTIMNGICPCORMDQY PHEHDMTPENFKCKDLRENYCRNPDS 360
 DB 301 ETTECIQGGEGYRGTVNTIMNGICPCORMDQY PHEHDMTPENFKCKDLRENYCRNPDS 360
 QY 361 ESPWCFETPDNIRVGYCSQIPNCDMSHGDCYRGNGKNYGNLSQTRSGLTCSMDKME 420
 DB 361 ESPWCFETPDNIRVGYCSQIPNCDMSHGDCYRGNGKNYGNLSQTRSGLTCSMDKME 420
 QY 421 DLHRHTEWEPDASKLENCRNPDDAHGWCCTGNPLIPMDYCPISRCGDTPTTVNL 480
 DB 421 DLHRHTEWEPDASKLENCRNPDDAHGWCCTGNPLIPMDYCPISRCGDTPTTVNL 480
 QY 481 DHPVISCARTKQLRVNGIPTRTNIGMNVSLARYRNKHICGSLIKESWVLTARQCPSPRD 540
 DB 481 DHPVISCARTKQLRVNGIPTRTNIGMNVSLARYRNKHICGSLIKESWVLTARQCPSPRD 540
 QY 541 LKDYEARMLGIDHVGKDEKCKOVLNVSQLVYBEGSDLVLMKLARPAVLDDFVSTIDLP 600
 DB 541 LKDYEARMLGIDHVGKDEKCKOVLNVSQLVYBEGSDLVLMKLARPAVLDDFVSTIDLP 600
 QY 601 NNGCTIPERTSCSYGMYTGLINVDGLLRVLAHLYTMNGECSOHHRGKVTLINESETCAG 660
 DB 601 NNGCTIPERTSCSYGMYTGLINVDGLLRVLAHLYTMNGECSOHHRGKVTLINESETCAG 660
 QY 661 AEKISGPEGDYGGLVCEQHMKRMVLAIVYGRGCAIPNRRGIFVRVAYAKVIHKII 720
 DB 661 AEKISGPEGDYGGLVCEQHMKRMVLAIVYGRGCAIPNRRGIFVRVAYAKVIHKII 720
 QY 721 LTKYKPOS 728
 DB 721 LTKYKPOS 728

RESULT 2

HGF_MOUSE STANDARD; PRT; 728 AA.
 ID HGF_MOUSE
 AC 008048; 064007; 061662;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hepatocyte growth factor precursor (Scatter factor) (SF)
 DE (Hepatopoietin A).
 DE HGF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 496-504.
 RC TISSUE=Mammary fibroblast;
 RX MEDLINE=94183257; PubMed=8135822;
 RA Sasaki M., Nishio M., Sasaki T., Enami J.;
 RT "Identification of mouse mammary fibroblast-derived mammary growth
 RT factor as hepatocyte growth factor.";
 RL Biochem. Biophys. Res. Commun. 199;772-779(1994).
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94363381; PubMed=8081873;
 RA Lee C.C., Kozak C.A., Yamada K.M.;
 RT "Structure, genetic mapping, and expression of the mouse Hgf/scatter
 RT factor gene.";
 RL Cell Adhes. Commun. 1:101-111(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94060105; PubMed=8241272;
 RA Liu Y., Michalopoulos G.K., Zarnegar R.;
 RT Molecular cloning and characterization of cDNA encoding mouse
 RT hepatocyte growth factor";
 RL Biochim. Biophys. Acta 1216:299-303(1993).
 CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
 CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS
 CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
 CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1- ALTERNATIVE PRODUCTS: A SHORT FORM OF HGF IS PRODUCED BY
 CC ALTERNATIVE RNA SPLICING. THE SEQUENCE SHOWN HERE IS THAT OF THE
 CC LONG FORM.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
 CC -1- SIMILARITY: BELONGS 4 KRINGLE FAMILY. ALSO KNOWN AS THE
 CC TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D10212; BAA01064.1; -;
 DR EMBL: D10213; BAA01065.1; -;
 DR EMBL: S71816; AAB31855.1; -;
 DR EMBL: X72307; CAA51054.1; ALT_INIT.
 DR HSSP: P14210; 1BHT.
 DR MGD: MGI:96079; Hgf.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00051; kringle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR SMART: SM00130; KR; 4.
 DR SMART: SM00473; PAN_AP; 1.
 DR SMART: SM00020; TYP_SPC; 1.
 DR PROSITE: PS00021; KRINGLE_1; 4.
 DR PROSITE: PS50070; KRINGLE_2; 4.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR Growth factor; Kringle; Glycoprotein; Serine protease homolog;
 KW Signal; Alternative splicing.
 FT CHAIN 1 32
 FT CHAIN 33 495
 FT CHAIN 496 728
 FT MOD_RES 33 33
 FT (BY SIMILARITY).
 FT HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
 FT HEPATOCYTE GROWTH FACTOR BETA CHAIN.
 FT PYROLIDONE CARBOXYLIC ACID
 FT (BY SIMILARITY).

```

FT DOMAIN 33 128 PAP.
FT DOMAIN 129 207 KRINGLE 1.
FT DOMAIN 212 289 KRINGLE 2.
FT DOMAIN 306 384 KRINGLE 3.
FT DOMAIN 392 470 KRINGLE 4.
FT DOMAIN 496 728 SERINE PROTEASE-LIKE.
FT DISULFID 71 97 BY SIMILARITY.
FT DISULFID 75 85 INTERCHAIN (BY SIMILARITY).
FT DISULFID 488 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 656 656 MISSING (IN SHORT ISOFORM).
FT VARSPLIC 163 167 N -> K (IN REF. 2).
FT CONFLICT 344 344 V -> L (IN REF. 2).
FT CONFLICT 479 479 R -> H (IN REF. 3).
FT CONFLICT 564 564
SQ SEQUENCE 728 AA: 82944 MW: A0381FC497534328 CRC64;

```

Query Match 92.28; Score 3804; DB 1; Length 728;

Best Local Similarity 90.88; Pred. No. 7.5e-277; Matches 660; Conservative 39; Mismatches 26; Indels 2; Gaps 1;

```

QY 1 MWYKLLPALLOHVLHLLHLLPFAIPYAEQORRRNTIHEFKSAKTTLIKIDPAKIK 60
  |||||
DB 2 MNGTKLLPVLLQHVLLHLLHLLHVAIPYAEQORRRNTIHEFKSAKTTLIKIDPAKIK 61
QY 61 TKKYNVADOCANRCTRNKGLPFTCKAFVFDKARKOCILPFPNNSSGYKKEFGHEFDLYE 120
  |||||
DB 62 TKKYNVADOCANRCTRNKGLPFTCKAFVFDKARKOCILPFPNNSSGYKKEFGHEFDLYE 121
QY 121 NKDIYRNCIIIGKSGSYGTYSITSGIKCOPMSMIPIHEHFLSSYRKGLOJENYCNRP 180
  |||||
DB 122 NKDIYRNCIIIGKSGSYGTYSITSGIKCOPMSMIPIHEHFLSSYRKGLOJENYCNRP 181
QY 181 RGEGBGWCFTSNPEVRYEVDIPQCESEVECMTCNGESYRGLMDHTESGKICQRMHQTP 240
  |||||
DB 182 RGEGBGWCFTSNPEVRYEVDIPQCESEVECMTCNGESYRGLMDHTESGKICQRMHQTP 241
QY 241 HRHGFLEPRIPDKFDONCYCNPDPGCRPMCYTLDPTRMETCYAIKTCADNTMDTVPL 300
  |||||
DB 242 HRHGFLEPRIPDKFDONCYCNPDPGCRPMCYTLDPTRMETCYAIKTCADNTMDTVPL 301
QY 301 ETTPCIGOGGSGYGTNTINGIIPCORMDSOYRPHEDMPENFKCKDLRNYCRNPDS 360
  |||||
DB 302 ETTPCIGOGGSGYGTNTINGIIPCORMDSOYRPHEDMPENFKCKDLRNYCRNPDS 361
QY 361 ESPWCFTTDPNIRVGYCSQIIPNCDMSHGQDCYRGNKNGKYNMGLSOTRSLGTCSMWDMKME 420
  |||||
DB 362 ESPWCFTTDPNIRVGYCSQIIPNCDMSHGQDCYRGNKNGKYNMGLSOTRSLGTCSMWDMKME 421
QY 421 DLHRIHTEWPDASKLNNYCNPNDDAHGPRCTYGNPLIPMDYCPISCEEDTPTTYNL 480
  |||||
DB 422 DLHRIHTEWPDASKLNNYCNPNDDAHGPRCTYGNPLIPMDYCPISCEEDTPTTYNL 481
QY 481 DHPVISCATKQQLVNVGIGIPRTINGMWVSLRYRNKNIHCGSLIKESVNLARACQFSPR 539
  |||||
DB 482 DHPVISCATKQQLVNVGIGIPRTINGMWVSLRYRNKNIHCGSLIKESVNLARACQFSPR 541
QY 540 -DLKDYEWAMLGIDHVGSGDEKCKQOVNLVSQLVYGPESGSDVLMLKLARPAVLDDFVSTID 598
  |||||
DB 542 KDLKDYEWAMLGIDHVGSGDEKCKQOVNLVSQLVYGPESGSDVLMLKLARPAVLDDFVSTID 601
QY 599 LPNKGCTIPREKTSYGVGWTGLINTDGLRVNHLVYIMGEEKSQHHRKGYVTLNESITC 658
  |||||
DB 602 LPNKGCTIPREKTSYGVGWTGLINTDGLRVNHLVYIMGEEKSQHHRKGYVTLNESITC 661
QY 659 AGAKKISGPGEGYGGGLVCEQHKMRVNLGVIYVGRGCAIPNPGIVVRAVYAKWIKH 718
  |||||
DB 662 AGAKKISGPGEGYGGGLVCEQHKMRVNLGVIYVGRGCAIPNPGIVVRAVYAKWIKH 721
QY 719 IILTYKV 725

```

```

DB 722 VILTYKL 728
  :|||||
RESULT 3
HGF_RAT ID HGF_RAT STANDARD; PRT; 728 AA.
AC P17945;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
DE (Hepatopoietin A).
GN HGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=90222197; PubMed=2139229;
RA Toshito K., Hagiya M., Nishizawa T., Seki T., Shimonishi M.,
RA Shimizu S., Nakamura T.;
RT "Reduced primary structure of rat hepatocyte growth factor and
  expression of the mRNA in rat tissues.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=91031482; PubMed=2146117;
RA Okajima A., Miyazawa K., Kitamura N.;
RT "Primary structure of rat hepatocyte growth factor and induction of
  its mRNA during liver regeneration following hepatic injury.";
RL Eur. J. Biochem. 193:375-381(1990).
CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATRE PARENCHYMAL
CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS
CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90102; BAA14133.1; -;
DR EMBL; X54400; CAA38266.1; -;
DR PIR; S13211; S13211.
DR PIR; A35644; A35644.
DR HSSP; P14210; 1BHT.
DR MEROPS; S01.978; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Trypsin.
DR Pfam; PF000051; kringle_4.
DR Pfam; PF000024; PAN_1.
DR Pfam; PF000089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_Ap_1.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.

```



```

DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRYP-Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
KW Kringle; Glycoprotein; Serine protease homolog; signal;
KW Polymorphism.
FT SIGNAL 1 18
FT CHAIN 19 711
FT DOMAIN 32 109
FT DOMAIN 110 186
FT DOMAIN 191 268
FT DOMAIN 283 361
FT DOMAIN 370 448
FT DOMAIN 484 711
FT DISULFID 56 78
FT DISULFID 60 66
FT DISULFID 110 186
FT DISULFID 131 169
FT DISULFID 157 181
FT DISULFID 191 268
FT DISULFID 194 324
FT DISULFID 212 251
FT DISULFID 240 263
FT DISULFID 283 361
FT DISULFID 304 343
FT DISULFID 332 355
FT DISULFID 370 448
FT DISULFID 391 431
FT DISULFID 419 443
FT DISULFID 468 588
FT DISULFID 507 523
FT DISULFID 602 667
FT DISULFID 632 646
FT DISULFID 657 685
FT CARBOHYD 72 72
FT CARBOHYD 296 296
FT CARBOHYD 615 615
FT VARIANT 13 13
FT VARIANT 13 13
FT VARIANT 212 212
FT VARIANT 212 212
FT CONFLICT 623 623
FT SEQUENCE 711 AA; 80379 MM; 596ED21F180290E4 CRC64;

```

Query Match 40.0%; Score 1652; DB 1; Length 711;
 Best Local Similarity 43.8%; Pred. No. 5,2e-116;
 Matches 317; Conservative 103; Mismatches 265; Indels 38; Gaps 14;

```

16 LHLHLPLIAPYAGQKKRNTIEFKSAKTLIKIDPAKIK--TKKNTADQCANR 73
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
6 LLLLTOLVLP---GQ---RSPILDQVLRGTLEQHLHLHVPQWQEDVADAECAGR 59
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
74 CTRKNGLPFTCKAVFDKARKQCLMFENSSSGVKKFEFGEFDYENKDYIRNCIIGK 133
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
60 C---GLMDRAFHYNVSSHCQQLPWTQHSPIRLARSGRDLFQKKDYVTCITMNG 115
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
134 RSYGVTSIKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCTSN 193
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
116 VGRGTMTATVGGGLPCQAMSHKFPMDHKYTPILRNG--LEENFCNPDGDPGGPGWCTTD 173
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
194 PEVYEVODIQCCSEVEEMTNGESYRGLMHTESGKICQKMDHOTPHKHKFLPERIYDK 253
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
174 PAVRFQSGIKSCREACVWCNGEYRGAVRTTSGRCQKMDLQHPQHFEPEPKFLDQ 233
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
254 GFQDNVCRNPDGGRPCWCTLDPTHTRMVEYCAIKTCADTMNDTVLPLETTTCIOGGEGY 313
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
234 GLDNYCRNPDGGRPCWCTTDPTHTRMVEYCAIKTCADTMNDTVLPLETTTCIOGGEGY 291
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 314 RGVNTIANGIPCOQRWDSQYPHEHDMTPENFKCKDLRENYCRNPDGSESPPGCTTDPNIR 373
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
DB 292 RGTANTTAGVYCOQRWDAIPIPHRTPEKYACKDLRENFCCRNPDSSEAPMCTTLRPGKR 351
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 374 VGYCQIENC-DMSGQDCYRGNGKNYGNLSQTRSGITCSMMDKNMEDLHRIIWEPPDA 432
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
DB 352 AAFVQYIRRCTDVPDRPODCYHAGGEYRGTSKTRGVCCQRMASAPTPHKPQFTFTSEPH 411
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 433 SKLENVYCRNPDGDAHGPWCYTGNPLIPMDYCPISGCEBDITPTIYNLDHP-----VISCA 488
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
DB 412 AOLNEFCRNPDSHGPWCYTMDPPTPDYCALRCADDPPSI--LDPPQVOYFEKKG 469
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 489 K-----TKQLRVNGIIPTRTNIGMWSLRYR--NKHICGSLKESVTLAROCFPS- 538
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
DB 470 KAVDRLDQRRSLRYVGGP--GNSPMYSLNRQGHGCGSLYKEQWILAROCFSSC 527
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 539 -RDLDYEAALGIDHVGGRDEKQVLYNSQLYVPEGSDVLYKLKRAVADDFVSTI 597
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
DB 528 HMLTNGYEWALGTLRPNQPHGEPSSLQRPVAKMCGSPSGQLVLKLESVTLNRVALI 587
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 598 DLPNGCTIPEKTSQSVYWGVTGLINQGLRVAHLIYMGKCSQHHRGKVTTLNESFI 657
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
DB 588 CLPPEWYVVPQTKCEIAGMGETKGTGNDTVLNLVINSQECNLIKRGV--RESEM 645
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 658 CAGAEKIGSGPECDYGGPGLVCEQHKRMRYLVGIYVGRCAIIPNRPPIVAVYAKWII 717
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
DB 646 CTEGLLAPYAGCEGYGGPLACFTINCVLBSIIIPNRYCANRMPAVTRVSVVDWII 705
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 718 KII 720
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :
DB 706 KVM 708
   ||||| : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 5
 HGFL_MOUSE
 ID HGFL_MOUSE STANDARD: PRT; 716 AA.
 AC P26928;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hepatocyte growth factor-like protein precursor (Macrophage stimulatory protein) (MSP).
 GN Mst1 OR HGFL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE=Liver;
 RX MEDLINE=92002017; PubMed=1832957;
 RA Friesner Degen S.J., Stuart L.A., Han S., Jamison C.S.;
 RT "Characterization of the mouse cDNA and gene coding for a hepatocyte growth factor-like protein: expression during development.";
 RL Biochemistry 30:9781-9791(1991).
 CC -!- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT CONSERVED.
 CC -!- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND ADRENAL.
 CC -!- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION. JUST BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS STABLE AFTERWARDS.
 CC -!- PM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE POLYPEPTIDES.
 CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPsin FAMILY. PLASMINOGEN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its

Db	60	C--GGLLDLCRAFNHYMMSSHGQQLPWHQSHLHQLHYHSSILCHLFLQKKDKDYVRICIMDNG	115
Qy	134	RSYGVGVITSKSGIKCOPMWSMTPIHEHSFLPSSYRGKIDQENYCRNPRGEEGGMPCSTSN	193
Db	116	VSYGVGVTAFTAGGILPCQAMSRFRPPNKKHKTTPRPKN--LEENFCNNPGDGRGPKFCYTTN	173
Qy	194	PEVYEVGCDLPCSEVYVCMTCNCESTRGLMDHTESKICQRRNDQTPRNHFLPRRYDK	253
Db	174	RSVFQSGCIKTCREAVCYLVLCNEDYRGEDVATESRECORDMLDQHPSHRPQPEKFDK	233
Qy	254	GFDNNYCRNPDGQRPCTYLLDHTEMEKYCAIKTCADN-----TMDTQVPLETTECSI	306
Db	234	DLKNTYCRNPDGSRPRMPCYTTDPNVERECODLPSCGPNLPPIYVKSCKSQRRKAKALNCF	293
Qy	307	QGGEGYRGTVNTIWNIGIPCOQRWDSQYPHENDHTPENFKCKDLRENYCRNPDGSESPCF	366
Db	294	RKGEGEDRGTTNNTSAGVPCQWRDADSPQHNRFVPEKYACKDLRENPGRNPDGSEAPMCF	353
Qy	367	TTDININVGICSDIPLNC-DMSHGQDCTRGNGKTKYGNLSQTSGLTCSMDKKNMEDLNRH	425
Db	354	TSRGLTSMACFOHLPCTCELVPEGGCYHSGEDYRGSVSKTKRGVQOCHN--SSETPPKR	411
Qy	426	IFWEPDASK--LLENYCRNPDGDAHRCMPCYTGPNLPIMQYCISSCEGPTPIYVNDHR	483
Db	412	QFTPTSPAQAGLEAFNFRNPDGDSHGRCYTTDPLDLYFCALDYCRDDQRPSTIDRPDQ	471
Qy	484	VI--SCA-----KTQKLVKVVNGIPIRTNIGMYSLYR-KNKIKGSGSLIKESVLLTAROCF	536
Db	472	VVEFKCGKRYDKSKMLKLVVGNR--GNSPMYVLSLRNQGQNHCGGSLVKEDQVLLTARCI	529
Qy	537	PS--RDLKDYBAMGLINDVNGRQDEKCKOYLVSQLVYRPESDILVLMKLAIRAVLDYV	594
Db	530	WSCHERPLTGEYMWLTGNQNPQGEANLQRPVAKAVCSGASQVLVLKLERYVLLMNV	589
Qy	595	STDLPNVCGIPIREKTSCTSYVGMGYGLINVOGLVALYLTMGCEKSONHNRKVLTLNE	654
Db	550	ALICLPDPQYVWPVGTICELTAGBESIGTSNNVLVAAMNYLSMQEBCVTKRGT--QE	647
Qy	655	SEICAGAEKIGSGCEDYDGGRLVCEONKIMRYLVGIYVGRSCATIPNRGIFVRVAYUAK	714
Db	648	SEITQGLVYVYACGCESDYGGRLACTTHOCWVLQGLIIPNRCAIPRYALGTRSVYVD	707
Qy	715	WIKKI 720	
Db	708	WINKM 713	
RESULT	6		
PLMN_PIG		STANDARD;	PRT; 790 AA.
AD	PO6867;		
IC	01-JAN-1988 (Rel. 06, Created)		
DT	01-FEB-1991 (Rel. 17, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Plasmidogen (EC 3.4.21.7).		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eultheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
RN	[1]		
RP	SEQUENCE OF 1-560.		
RA	Schaller J., Martl T., Roesseler S.J., Kaempfer U., Rickli E.E.;		
RT	"Amino acid sequence of the heavy chain of porcine uro-		
RT	of the carbohydrate attachment sites with the human and bovine		
RT	species".		
RL	Fibrinolysis 1:91-102(1987).		
RN	[2]		
RP	SEQUENCE OF 450-790.		
RX	MEDLINE=85203907; PubMed=3846533;		
RA	Martl T., Schaller J., Rickli E.E.;		
RT	"Determination of the complete amino-acid sequence of porcine		
RT	miniplasminogen".		
RL	Eur. J. Biochem. 149:279-285(1985).		

[3]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=88185329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 RT Gerold G.J., van Halbeek H., Vilgertshart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 RT plasmidogen. Species specificity in relation to sialylation and
 RT fucosylation patterns.";
 RL Eur. J. Biochem. 173:57-63(1988).
 CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 CC AND INFLAMMATION. IN OVULATION IT WEAKENS THE WALLS OF THE
 CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -1- PTM: N-LINKED GLYCAN CONTAINS N-ACETYLACTOSAMINE, SIALIC ACID AND
 CC IS CORE FUCOSYLATED. O-LINKED GLYCANS CONSIST OF GAL-GALNAc
 CC DISACCHARIDE WITH IS MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES
 CC (MICROHETEROGENEITY).
 CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ALSO KNOWN AS THE
 CC TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
 CC PIR: A25834; A25834.
 CC PIR: S03733; S03733.
 CC HSSP: P00747; SHPG.
 CC MEROPS: S01.233; -.
 CC GlycoSuiteDB: P06867; -.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00051; Kringle; 5.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; Trypsin; 1.
 DR PRINTS: PRO0022; CHYMOTRYPSIN.
 DR PRINTS: PRO0018; KRINGLE.
 DR SMART: SM00130; KR; 5.
 DR SMART: SM00473; PAN_APP; 1.
 DR SMART: SM00020; TRYP_SPEC; 1.
 DR PROSITE: PS00021; KRINGLE_1; 5.
 DR PROSITE: PS50070; KRINGLE_2; 5.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; FALSE_NEG.
 DR PROSITE: PS00133; TRYPSIN_SER; 1.
 KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.
 FT CHAIN 1 560 PLASMIN HEAVY CHAIN A.
 FT CHAIN 561 790 PLASMIN LIGHT CHAIN B.
 FT DOMAIN 561 790 SERINE PROTEASE.
 FT DOMAIN 84 162 KRINGLE 1.
 FT DOMAIN 166 243 KRINGLE 2.
 FT DOMAIN 256 333 KRINGLE 3.
 FT DOMAIN 358 435 KRINGLE 4.
 FT DOMAIN 461 540 KRINGLE 5.
 FT ACT_SITE 602 602 CHARGE RELAY SYSTEM.
 FT ACT_SITE 645 645 CHARGE RELAY SYSTEM.
 FT ACT_SITE 740 740 CHARGE RELAY SYSTEM.
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 289 289 /FTID-CAR_000019.
 FT CARBOHYD 340 340 O-LINKED (GALNAc. . .).
 FT CARBOHYD 340 340 /FTID-CAR_000020.
 SO SEQUENCE 790 AA; 88592 MW; F04EA0674BCD58E CRC64;

Query Match

35.1%; Score 1446.5; DB 1; Length 790;

Best Local Similarity 36.9%; Pred. No. 1.3e-100;
 Matches 294; Conservative 114; Mismatches 232; Indels 157; Gaps 20;
 QY 56 ALKIKTKVNTADQCANRCKTNRKGLPEFTGKAFVFPKARKOCPLFPNSMSGVKKFEFHE 115
 DB 16 SLSRKQVAAKRSVECAKAEALTN--FICRAQYISKDQOCVVAEDNSKTSPIAHM--RD 71
 QY 116 FDLVNEKDIYRNCITIGKSGYVSTKSGITKQCPWSSMIPHEHSFLPSYRGKDLQEN 175
 DB 72 VVLFERKRLISLCKTGNGKNGRYGTSKTSKVICQKWSVSPHIKYSPEKFLAGLDEN 131
 QY 176 YCRNPRGEGSGWCTSNPEVEYECDIPOCESEVCHNCNCSYGLMDHESGTCQW 235
 DB 132 YCRNPNDKSGWCTYTPTEFDFCDIPEC-EDSCMCSGHEGKISKTSKTSWTECOW 190
 QY 236 DHQTRHNEFLPERYPDKGFDNNYCRNPDGQRPMPCTYLDPTREYCAITCADNTMD 295
 DB 191 GSQSPHAGYLPKRPNNLKNYCRNPDGERPRPCFTTDPKRRKREFCDIRC--TIIPP 248
 QY 296 TDVPLETECLOGGEGYRGYVNTWNGIPQORWDQYPEHEDMTPENFKCKDLRENYCR 355
 DB 249 TSGP--TYQCLKRGENTRGVSYASGHTCQWASQSPHKNHRTPENFKCKDLRENYCR 306
 QY 356 NEDGSESEPCFTTDPENIRGVCSQIPNCDMS--H-----GDDCYRGNGKN 398
 DB 307 NPDGETAPWCYTTTDESEVWDIC-KIPSGSSTTTEHLDAPVPEQTPVADDCYRGNGES 365
 QY 399 YMGNISQTRSGITCSM-----DK-----417
 DB 366 YRGTSTTITGRKCSQSVSMTPHREKTPGNFPNAGLTMNCRNPDADKSWCTTDPV 425
 QY 418 -----NMEDL-----422
 DB 426 RWEYCNLKKCSETEQVTNFPALIAQVPSVEDLSDECMFGNRRYRGKRTTVAGVPCQEW 485
 QY 423 -----HRIIEWEPDASK---LNENYCRNPDADAGHPCYTNPLIPDYCPISRCEDDT 474
 DB 486 AAOEPHRSITPETNPRAGLEKNYCRNPDADDGMPCTTNPCKLTDYCDVPCVYSS- 544
 QY 475 PTYNLDHPVYSCAKTK-----OLRVYNG-IPRTNIGMWSLRYRNK-HICGSLIKE 526
 DB 545 -----FDCKRPKEPKCPARVYGGCVSIPHSMPQDISLRYRGHFGCGTLISP 594
 QY 527 SWVLTAROCF-PSRDLMDYEMLGIDVHGKDEKCOVLNVQVYVGPESGDLVLMKLA 585
 DB 595 EMTLAKHLEKSSPSKYITLGAHEHYLGE--VQEDVDSLFKEPSEADIALKLIS 652
 QY 586 RPAVLDFVSTIDLPNGCTIPEKTSQVYGMGTGLINVDGLRVAHLVIMGENKCSOH 645
 DB 653 SPAVITDKVIPACLPPTNPNVADRTACTYITGWEETKGTGAGLLEKRLVLENKYNRY 712
 QY 646 H--RGKTLINSELCACAEKIGSPCEGVDGGLYVCSQHMKNRYLYIVYRGCAIPNRP 703
 DB 713 EYLGKVSFN--ELCAGHLAGGIDSCGDSGLVCEPKDKYILQGVTSWGLCALPNKP 770
 QY 704 GIEVRAVAVYAKMIKII 720
 DB 771 GVVYRVSRTVWIEIM 787
 RESULT 7
 ID PLAN_BOVIN STANDARD; PRT; 812 AA.
 AC P06868; 028162;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Plasmidogen precursor (EC 3.4.21.7).
 GN PLG.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.


```

Db      510  A A E R H O S I T P E R I N P Q S G L E R N T C R N P D D V A N G P M C Y T I N N P K F P D Y C V P C E S S -- 567
Qy      475  P T I V N L D H P V I S C A K T K ----- Q L R V V N G I P T P T N I - G M V N S L Y R N K I I G C S L I K E S 527
Db      568  ----- F D C G K P R K E P K C S G R I A I G C V S R P H S M P M O V S L R S R S R H P C G T L I S P K 617
Qy      528  W L T A R O C F P S - R D I K O E A M L G I H D V I N G R D E K C Q V I A N S O L V Y P E G S D L Y M K L A R 588
Db      618  W L T A R H O L D I N I L A S F K V I L G A H N E K V R - E O S V O E I P V S R F R E P S O A D I A L L T L S R 679
Qy      587  P A V I D D F S T I D L P N Y C G T I P E K T S C S V Y G M G Y G L I N Y O D G L P V A N H Y I T G N E K S O H N 646
Db      676  P A I T K E V I P A C L R P P N N M V A R I E C Y T I T G E Q I G F G E S L L E A H L P Y T E N K V C R N E 733
Qy      647  R G K Y T L N E S I E C A G A E K I G S G P C E G D Y G R I P C E O H K M R M Y I V B P G C A C I P N R G I F 706
Db      736  Y L D G V K R P T E L C A C H L I G T D S C Q D S G R P L V C F E K D Y I L O G V T S N G L G C A R P N K R G V Y 795
Qy      707  V R V A Y A K M I I K 718
       796  V R V S P Y V F W I E E 807

```

RESULT	8
PLMN_MOUSE	
ID	PLMN_MOUSE
AC	STANDARD:
AC	P20918;
DT	01-FEB-1991 (Rel. 17, Created)
DT	01-FEB-1991 (Rel. 17, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Plasminogen precursor (EC 3.4.21.7) [contains: Angiotatin].
GN	PLG.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxId=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=91184812; PubMed=2081600;
RX	Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
RT	"Characterization of the cdna coding for mouse plasminogen and
RT	localization of the gene to mouse chromosome 17.";
RL	Genomics 8:49-61(1990).
RN	[2]
RP	CHARACTERIZATION OF ANGIOTATIN, AND PARTIAL SEQUENCE.
RX	MEDLINE=95042728; PubMed=752077;
RA	O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA	Moses M., Lane W.S., Cao Y., Saxe E.H., Polman J.;
RT	"Angiotatin: a novel angiogenesis inhibitor that mediates the
RT	suppression of metastases by a Lewis lung carcinoma.";
RL	Cell 79:315-328(1994).
CC	-1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC	A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC	EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC	AND INFLAMMATION. IN OVULATION IT WEAKENS THE WALLS OF THE
CC	GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC	ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC	AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC	LAMININ AND VON WILLEBRAND FACTOR.
CC	-1- FUNCTION: ANGIOTATIN IS AN ANGIOGENESIS INHIBITOR THAT BLOCKS
CC	NEOVASCULARIZATION AND GROWTH OF EXPERIMENTAL PRIMARY AND
CC	METASTATIC TUMORS IN VIVO.
CC	-1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC	ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC	FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC	-1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC	IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC	-1- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC	INVOLVES ONLY CLEAVAGE AFTER ANG-581, RESULTING IN 2 CHAINS HELD
CC	TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC	ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC	-1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.

CC	-1	SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC		TRYPsin FAMILY PLASMINOGEN SUBFAMILY.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC		the European Bioinformatics Institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC		or send an email to license@isb-sib.ch).
CC		-----
DR	CC	EMBL; J04766; AAA50168.1; -
DR	DR	PIR; A38514; A38514.
DR	DR	HSSP; P00747; IPMK.
DR	DR	MEROPS; S01.233; -
DR	DR	MGD; MGI:97620; Plg
DR	DR	InterPro; IPR001314; Chymotrypsin.
DR	DR	InterPro; IPR000001; Kringle.
DR	DR	InterPro; IPR003014; PAN.
DR	DR	InterPro; IPR003609; Pan_app.
DR	DR	InterPro; IPR001254; Trypsin.
DR	DR	Pfam; PF00051; Kringle5.
DR	DR	Pfam; PF00024; PAN.1.
DR	DR	Pfam; PF00089; trypsin.1
DR	DR	PRINTS; PR00722; CHYMOTRYPsin.
DR	DR	PRINTS; PR00018; KRINGLE.
DR	DR	SMART; SMO0130; KR. 5.
DR	DR	SMART; SMO0473; PAN_App.1.
DR	DR	SMART; SMO0020; Tryp_Spc.1.
DR	DR	PROSITE; PS000021; KRINGLE.1; 4.
DR	DR	PROSITE; PS50070; KRINGLE.2; 5.
DR	DR	PROSITE; PS50240; TRYPSIN_DOM.1.
DR	DR	PROSITE; PS00134; TRYPSIN_HIS.1.
DR	DR	PROSITE; PS00135; TRYPSIN_SER.1.
KW		Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW		Tissue remodeling; Blood coagulation; Kringle; zymogen; signal.
FT	FT	SIGNAL 1 19
FT	FT	CHAIN 20 812 PLASMINOGEN.
FT	FT	CHAIN 20 581 PLASMIN HEAVY CHAIN A.
FT	FT	PEPTIDE 20 97 ACTIVATION PEPTIDE.
FT	FT	CHAIN 98 581 PLASMIN SHORT FORM OF CHAIN A.
FT	FT	CHAIN 98 7436 ANGIOSTATIN.
FT	FT	CHAIN 103 812 PLASMIN LIGHT CHAIN B.
FT	FT	DOMAIN 103 181 KRINGLE 1.
FT	FT	DOMAIN 184 262 KRINGLE 2.
FT	FT	DOMAIN 184 352 KRINGLE 3.
FT	FT	DOMAIN 377 454 KRINGLE 4.
FT	FT	DOMAIN 377 454 KRINGLE 5.
FT	FT	DOMAIN 481 812 SERINE PROTEASE.
FT	FT	DOMAIN 582 812 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	FT	ACT_SITE 624 624 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	FT	ACT_SITE 667 667 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	FT	ACT_SITE 762 762 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	FT	DISULEFID 49 73 BY SIMILARITY.
FT	FT	DISULEFID 53 61 BY SIMILARITY.
FT	FT	DISULEFID 103 181 BY SIMILARITY.
FT	FT	DISULEFID 124 164 BY SIMILARITY.
FT	FT	DISULEFID 152 176 BY SIMILARITY.
FT	FT	DISULEFID 185 262 BY SIMILARITY.
FT	FT	DISULEFID 188 316 BY SIMILARITY.
FT	FT	DISULEFID 206 245 BY SIMILARITY.
FT	FT	DISULEFID 234 257 BY SIMILARITY.
FT	FT	DISULEFID 275 352 BY SIMILARITY.
FT	FT	DISULEFID 296 335 BY SIMILARITY.
FT	FT	DISULEFID 324 347 BY SIMILARITY.
FT	FT	DISULEFID 377 454 BY SIMILARITY.
FT	FT	DISULEFID 398 437 BY SIMILARITY.
FT	FT	DISULEFID 426 449 BY SIMILARITY.
FT	FT	DISULEFID 481 560 BY SIMILARITY.
FT	FT	DISULEFID 502 543 BY SIMILARITY.
FT	FT	DISULEFID 531 555 BY SIMILARITY.
FT	FT	DISULEFID 568 687 BY SIMILARITY.
FT	FT	DISULEFID 578 586 INTERCHAIN (BY SIMILARITY).
FT	FT	DISULEFID 609 625 INTERCHAIN (BY SIMILARITY).

FT	DISULFID	701	768	BY SIMILARITY.
FT	DISULFID	731	747	BY SIMILARITY.
FT	DISULFID	758	786	BY SIMILARITY.
SQ	SEQUENCE	812 AA;	90846 MM;	D3A474A4FC2256F8 CRC64;

Query Match	33.4%	Score 1379.5;	DB 1;	length 812;
Best Local Similarity	34.4%	Pred. No. 1.4e-95;		
Matches 288;	Conservative 119;	Mismatches 262;	Indels 167;	Gaps 22

QY	15	VLLHLLPLPAIPAEEOORRRNTHFFKSAKTLTKIDPALKIKRKVNTA----	DOCA	71
Db	7	ILFLFLLL-----KPO-----GSLDGYISTOGASLFL-----	TKOLAAGVSDCL	50
QY	72	NCRTRNKGLPPTCKAFVDEARAKOCIMFPPNSMSSGVKKEFGHEFDLYENKDYIRNCIIG		131
Db	51	AKCEGE--TDFVCSFQYHSKEOCCAIMAESKSSIIIRM--RDVILFEKRYVLSCEKTG		106
QY	132	KGREYKCTVSTITSIGICQPMSSMIIPHEHSFLPSYKGLDQENYCNPNPGESEGPCFT		191
Db	107	IGNRYRTMSTRKTSGVACQKMGAFEPHPVNPSPSTHNEGLEENYCNPNDEGPCYCT		166
QY	192	SNEPEYEVCDIPQCSSEVEMCTNGESYRGIMDTEBEGKICQMRDQDTPRHAKLPERPY		251
Db	167	TDPKRRDYCNIPCEE--ECMYCSGEKYEKISKTMSGDLQOAMDSPPAHGVIAPKPF		222
QY	252	DKEFDNYCNPNDOGPPRWCYTTDPPHTRWEXCAIKTCADNTMNDVPLETTECIGOGGE		311
Db	226	SKNLKNMYCNHPDQEPFRPWCFTTDPTRKWECDIPRCT-----PPPPSPPLYOCLKRGGE		281
QY	312	GYRGVTVIWNNGIPCOHWSQYRPHENMTPENFCKDLRENYCNPNCGSPMPCFTDPN		371
Db	282	NYRGTASVTVYSGTKQWMSQEPHRRHRTPEPFCKMLEENYCNPNPGETRAPWCYTTDSQ		341
QY	372	IRVQCSQIPNCDMSHS-----QDQYRGNGKNYMNQNLQSTRSGLTCSM		414
Db	342	LRWEYC-EIPSCSSASAPDQSDSVPRPEQTPYVQECYQSDQSGSRGTSSTITIGKKQCS		400
QY	415	WDRKMEDLHRHIEF---PDASKINENYCNPNDDAHGPMYCIGNLIPMDYCPISRCBG		471
Db	401	WAAMFP--HRHSKTPENFPDAG--LENNYCNPNDDG--KGPMCYTTPDSVRWEYCNLKRCS		455
QY	472	DTPTIYNLHPVYS-----		486
Db	456	ETGSGVEL--PYVSOEPPSGPDSSETDCMANGKDYRGKTAVTAAAGTPOCGMAAOEPRHR		513
QY	487	-----CAKTKOL-----		493
Db	514	SIFTPQNPRADEKKNYCNPNDDGVNCPWCYTTNPRKLYDCDIPCLASASSFEGCAQY		573
QY	494	-----RYVNG-IPRTINIGMWSLKR--NKHICGSLIKESWULTAROCFSPRLKD		543
Db	574	EPKKCPBRYVGGCVAANDHSPMWOISLRTFTGQFCCGTLIAPBMYVLAANCHESSRPE		633
QY	544	-YEAMLCIHPVHGKRGDECKOVLNVSOLVYGPBESDVLVIMKLAPAYLDFVSTIDLPNY		602
Db	634	FYKVIITLCAHEYIRGLD--VOEISVALILIEPNRNDTALLKLSRPAITTTBKVIAPCLPSP		691
QY	603	GCTIPEKTSVYGMVGTGLINDGLIRVAHLIYMGENKCSQHHRGKVTULNESEICAGAE		663
Db	692	NYMADRTICYITGMGETOGTFGAGRLKEAQLPIYENKVCNRYEYLLNNRKYSTELACQL		751
QY	663	KISGPGCBGYGRLVDEQHKMRKRVGLVYIPGRCAIIPNPGIFVRYRAYAKWTHK		718
Db	752	AGGVDSCQGBSGGGLVCFEDKTYILQGVSTSMGACAPRNKPGVYVARSFLVDWIER		807

RESULT	9
PLMN_HUMAN	
ID	STANDARD; 810 AA
PLMN_HUMAN	
AC	P00747;
DT	21-JUL-1986 (Rel. 01, Created)
DT	01-MAR-1989 (Rel. 10, Last sequence update)

DE 16-OCT2001 (Rel. 40, Last annotation update)
 DE plasmnogen precursor (EC 3.4.21.7) [contains: Angiostatin].
 GN pig.
 OC Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90202879; PubMed=2318848;
 RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
 RT "Characterization of the gene for human plasminogen, a key proenzyme
 RT in the fibrinolytic system.";
 RL J. Biol. Chem. 265:6104-6111(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=07162490; PubMed=3030813;
 RA Forsgren M., Røden B., Israelsson M., Larsson K., Heden L.-O.;
 RT "Molecular cloning and characterization of a full-length cDNA clone
 RT for human plasminogen.";
 RL FEBS Lett. 213:254-260(1987).
 RN [3]
 RP SEQUENCE OF 20-810.
 RA Sottrop-Jensen L., Petersen T.E., Magnusson S.;
 RL Submitted (JUL-1977) to the PIR data bank.
 RN [4]
 RP SEQUENCE OF 292-810 FROM N.A.
 RX MEDLINE=05023311; PubMed=6148961;
 RA Malinowski D.P., Sadler J.E., Davie E.W.;
 RT "Characterization of a complementary deoxyribonucleic acid coding for
 RT human and bovine plasminogen.";
 RL Biochemistry 23:4243-4250(1984).
 RN [5]
 RP SEQUENCE OF 20-100.
 RX MEDLINE=75093329; PubMed=122932;
 RA Wiman B., Wallen P.;
 RT "Structural relationship between 'glutamic acid' and 'lysine' forms
 RT of human plasminogen and their interaction with the NH2-terminal
 RT activation peptide as studied by affinity chromatography.";
 RL Eur. J. Biochem. 50:489-494(1975).
 RN [6]
 RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
 RA Sottrop-Jensen L., Cleys H., Zajdel M., Petersen T.E., Magnusson S.;
 RL (in) Davidsson J.F., Rowan R.W., Samama M.M., Desnoyers P.C. (eds.)
 RL Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
 RL Raven Press, New York (1978).
 RN [7]
 RP SEQUENCE OF 483-604.
 RX MEDLINE=76043692; PubMed=126863;
 RA Wiman B., Wallen P.;
 RT "Amino-acid sequence of the cyanogen-bromide fragment from human
 RT plasminogen that forms the linkage between the plasmin chains.";
 RL Eur. J. Biochem. 58:539-547(1975).
 RN [8]
 RP SEQUENCE OF 581-810.
 RX MEDLINE=77225245; PubMed=142009;
 RA Wiman B.;
 RT "Primary structure of the B-chain of human plasmin.";
 RL Eur. J. Biochem. 76:129-137(1977).
 RN [9]
 RP ACTIVE SITE.
 RX MEDLINE=73149248; PubMed=4694729;
 RA Robbins K.C., Bernabe P., Arradon L., Summaria L.;
 RT "The primary structure of human plasminogen. II. The histidine loop
 RT of human plasmin: light (B) chain active center histidine sequence.";
 RL J. Biol. Chem. 248:1631-1633(1973).
 RN [10]
 RP ACTIVE SITE.
 RX MEDLINE=69234739; PubMed=4240117;
 RA Groskopf W.R., Summaria L., Robbins K.C.;
 RT "Studies on the active center of human plasmin. Partial amino acid
 RT sequence of a peptide containing the active center serine residue.";
 RL J. Biol. Chem. 244:3590-3597(1969).

RP [11]
 RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
 RX MEDLINE-82213905; PubMed-6919539;
 RA Trexler M., Vail Z., Patchy L.;
 RT "Structure of the omega-aminocarboxylic acid-binding sites of human
 RT plasminogen. Arginine 70 and aspartic acid 56 are essential for
 RT binding of ligand by kringle 4 ";
 RL J. Biol. Chem. 257:7401-7406(1982).
 RN [12]
 RN FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.
 RX MEDLINE-85054794; PubMed-6095426;
 RA Vail Z., Patchy L.;
 RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34
 RT are essential for fibrin affinity of the kringle 1 domain.";
 RL J. Biol. Chem. 259:13690-13694(1984).
 RN [13]
 RN PHOSPHORYLATION SITE SER-597.
 RP MEDLINE-97345939; PubMed-9201958;
 RX Wang H., Proctor M., Brethauer R.K., Castellino F.J.;
 RA "Serine-578 is a major phosphorylation locus in human plasma
 RT plasminogen.";
 RL Biochemistry 36:8100-8106(1997).
 RN [14]
 RN CARBOHYDRATE-LINKAGE SITES.
 RP MEDLINE-88185329; PubMed-3356193;
 RX Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 RA Gerwig G.J., van Halbeek H., Vliegenhart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 RT plasminogen. Species specificity in relation to sialylation and
 RT fucoylation patterns.";
 RL Eur. J. Biochem. 173:57-63(1988).
 RN [15]
 RN CARBOHYDRATE-LINKAGE SITE 268.
 RP MEDLINE-97207306; PubMed-9054441;
 RX Plite-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,
 RA Plizo S.V.;
 RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
 RT human plasminogen 2.";
 RL J. Biol. Chem. 272:7408-7411(1997).
 RN [16]
 RN CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
 RP MEDLINE-95042728; PubMed-7525077;
 RX O'Reilly M.S., Holmgren L., Shung Y., Chen C., Rosenthal R.A.,
 RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
 RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
 RT suppression of metastases by a Lewis lung carcinoma.";
 RL Cell 79:315-328(1994).
 RN [17]
 RN CHARACTERIZATION OF ANGIOSTATIN.
 RP MEDLINE-9728710; PubMed-9102221;
 RX Slin B.K., O'Reilly M.S., Liang H., Fortler A.H., He W., Madsen J.W.,
 RA Lapcevich R., Nacy C.A.;
 RT "A recombinant human angiotensin protein inhibits experimental primary
 RT and metastatic cancer.";
 RL Cancer Res. 57:1329-1334(1997).
 RN [18]
 RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
 RP MEDLINE-92031502; PubMed-1657148;
 RX Moolichak A.M., Tullinsky A., Ravichandran K.G.;
 RA "Crystal and molecular structure of human plasminogen kringle 4
 RT refined at 1.9-A resolution.";
 RL Biochemistry 30:10576-10588(1991).
 RN [19]
 RN X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
 RP MEDLINE-92031503; PubMed-1657149;
 RX Wu T.-P., Padmanabhan K., Tullinsky A., Moolichak A.M.;
 RA "The refined structure of the epsilon-aminocaproic acid complex of
 RT human plasminogen kringle 4.";
 RL Biochemistry 30:10589-10594(1991).
 RN [20]
 RN X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
 RP Stec B., Yamano A., Whitlow M., Teeter M.W.;
 RA Submitted (JUN-1995) to the PDB data bank.

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2002, 17:45:54 ; Search time 34.17 Seconds
(without alignments)
3685.658 Million cell updates/sec

Title: US-09-423-516-2

Perfect score: 4126
Sequence: 1 MMVTKLPALLQHVHLHL.....VAYYAKMKIKILTKVPS 728

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp_unclassified:*
15: sp-virus:*
16: sp_bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3880	94.0	728	6 Q9BH09	Q9BH09 felis silve
2	3170	76.8	726	13 Q90978	Q90978 gallus gall
3	2806	68.0	710	13 Q91402	Q91402 xenopus lae
4	1704	41.3	716	13 Q91691	Q91691 xenopus lae
5	1652.5	40.1	717	13 P70006	P70006 xenopus lae
6	1632	39.6	290	4 Q02935	Q02935 homo sapien
7	1627	39.4	704	13 Q90865	Q90865 gallus gall
8	1624	39.4	296	4 Q14519	Q14519 homo sapien
9	1615	39.1	716	11 P70521	P70521 rattus norv
10	1605	38.9	716	11 Q91XG8	Q91XG8 mus musculu
11	1579	38.3	709	13 Q90ZM6	Q90ZM6 brachydanto
12	1452	35.2	806	6 Q18783	Q18783 macropus eu
13	1401.5	34.0	812	11 Q9R0W3	Q9R0W3 rattus norv
14	1388.5	33.7	812	11 Q91WJ5	Q91WJ5 mus musculu
15	1371	33.2	810	4 Q15146	Q15146 homo sapien
16	1365	33.1	567	4 Q13208	Q13208 homo sapien

17	1364.5	33.1	648	4 Q9H1V4	Q9H1V4 homo sapien
18	1139	27.6	208	4 Q9BYM0	Q9BYM0 homo sapien
19	1139	27.6	210	4 Q13494	Q13494 homo sapien
20	1040	25.2	215	13 Q42341	Q42341 gallus gall
21	1036	25.1	211	11 Q55027	Q55027 mus musculu
22	945	22.9	2869	6 Q28398	Q28398 erinaceus e
23	785	19.0	454	6 Q46506	Q46506 papio hamad
24	654.5	15.9	313	13 Q9PU78	Q9PU78 crocodylus
25	615	14.9	109	6 Q9N1B8	Q9N1B8 ovis aries
26	556.5	13.5	334	6 Q46507	Q46507 papio hamad
27	555.5	13.5	286	4 Q13209	Q13209 homo sapien
28	549.5	13.3	559	11 Q91VP2	Q91VP2 mus musculu
29	523.5	12.7	516	4 Q98U99	Q98U99 homo sapien
30	500	12.1	607	13 Q91001	Q91001 gallus gall
31	494.5	12.0	608	13 Q9PTW7	Q9PTW7 struthio ca
32	471	11.4	597	11 Q35727	Q35727 mus musculu
33	462	11.2	616	6 Q97507	Q97507 sus scrofa
34	439	10.6	75	6 Q9BGN9	Q9BGN9 bos taurus
35	435.5	10.6	385	5 Q25101	Q25101 herdamia m
36	435	10.5	265	13 Q90WS3	Q90WS3 elaphe sp.
37	423.5	10.3	560	4 Q14520	Q14520 homo sapien
38	402	9.7	395	4 Q9BZM1	Q9BZM1 homo sapien
39	393	9.5	431	4 Q969W6	Q969W6 homo sapien
40	365	8.6	273	11 Q921N4	Q921N4 mus musculu
41	354	8.6	339	11 Q99L44	Q99L44 mus musculu
42	344.5	8.3	389	13 Q9PVX7	Q9PVX7 xenopus lae
43	340.5	8.3	643	6 Q97506	Q97506 sus scrofa
44	338.5	8.2	329	13 Q42272	Q42272 xenopus lae
45	338	8.2	275	4 Q96RZ6	Q96RZ6 homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	728 AA.
Q9BH09	Q9BH09			
AC	Q9BH09			
DT	01-JUN-2001 (TREMREL. 17, Created)			
DT	01-JUN-2001 (TREMREL. 17, Last sequence update)			
DT	01-DEC-2001 (TREMREL. 19, Last annotation update)			
DE	HEPATOCTE GROWTH FACTOR HGF.			
OS	Felis silvestris catus (Cat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.			
OX	NCBI_TaxID=9685;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LIVER;			
RA	Kobayashi Y., Nakamura N., Ishizaka T., Masuda K., Ohno K.,			
RA	Tsujiyama H.;			
RT	"Molecular cloning of feline hepatocyte growth factor (HGF) cDNA.";			
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.			
DR	EMBL: AB046610; BAB21499.1; -			
DR	HSSP: P14210; 1BHT			
DR	InterPro: IPR001314; Chymotrypsin.			
DR	InterPro: IPR000001; Kringle.			
DR	InterPro: IPR003014; PAN.			
DR	InterPro: IPR003609; Pan_app.			
DR	InterPro: IPR001254; Trypsin.			
DR	Pfam: PF00051; kringle; 1.			
DR	Pfam: PF00024; PAN; 1.			
DR	Pfam: PF00069; trypsin; 1.			
DR	PRINTS: PR00722; CHYMOTRYPSIN.			
DR	PRINTS: PR00018; KRINGLE.			
DR	SMART: SM00130; KR; 4.			
DR	SMART: SM00473; PAN_AP; 1.			
DR	SMART: SM00020; TRYP_SPE; 1.			
DR	PROSITE: PS00021; KRINGLE_1; UNKNOWN_4.			
DR	PROSITE: PS50070; KRINGLE_2; 4.			
DR	PROSITE: PS50240; TRYPSIN_DOM; 1.			

KW Hydrolase; Serine protease.
SQ SEQUENCE 728 AA: 83067 MW: 807F4A333D1E190A CRC64;

Query Match 94.0%; Score 3880; DB 6; Length 728;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 660; Conservative 28; Mismatches 18; Indels 4; Gaps 2;

QY 1 MMTVTLPLALLQVHLHLHLPIAIPYAEGRKRRNTHERKSAKTLIKIDPALKIK 60
DB 1 MMTVTLPLVLLQVHLHLHLHLPIAIPYAEGRKRRNTHERKSAKTLIKIDPALKIK 58
QY 61 TTKVNTADOCANCRNRKNGLPFTCAKAFVDFKARQCLMFPNMSGGKKKEGHEFDLYE 120
DB 59 TTKMTADOCANCRNRKNGLPFTCAKAFVDFKARQCLMFPNMSGGKKKEGHEFDLYE 118
QY 121 NKDYIRNCIIIGKRSYKGTASTITSGIKCOPWSSMTIHEHSFLPSSYRGKDLQENYCRNP 180
DB 119 NKDYIRNCIIIGKRSYKGTASTITSGIKCOPWSSMTIHEHSFLPSSYRGKDLQENYCRNP 178
QY 181 RGEEGPWCFTSNPEVREYVCDIPQCESEVEMTCNGESYRGMDHESGKICQRMHQTP 240
DB 179 RGEEGPWCFTSNPEVREYVCDIPQCESEVEMTCNGESYRGMDHESGKICQRMHQTP 238
QY 241 HRHKEFLPERYPDKGFDDNYCRNPDPGPRWCYTLDPTHEWYCAIKTCADNTMNDTDPVL 300
DB 239 HRHKEFLPERYPDKGFDDNYCRNPDPGPRWCYTLDPTHEWYCAIKTCADNTMNDTDPVL 298
QY 301 ETTECTIOGEGRGYVNTIWNIGIPQCRMDSQYPHEHDMTPENFKCKDLRENYCRNPDS 360
DB 299 ETTECTIOGEGRGYVNTIWNIGIPQCRMDSQYPHEHDMTPENFKCKDLRENYCRNPDS 358
QY 361 ESPMCEFTDPNINRGYCSOIPNCDMSHGDDCYRGKSKNTMGLSOTRSGILCSMDKME 420
DB 359 ESPMCEFTDPNINRGYCSOIPNCDMSHGDDCYRGKSKNTMGLSOTRSGILCSMDKME 418
QY 421 DLHRHTEMPEPDSKLNENYCRNPDDDAHGPCMTCYTNPLIPMDYCPISCEEDTPTIYNL 480
DB 419 DLHRHTEMPEPDSKLNENYCRNPDDDAHGPCMTCYTNPLIPMDYCPISCEEDTPTIYNL 478
QY 481 DHPVISCATKQLRVNGCIPTFTNIGMMVSLRYRKHICGSLIKESVYLTAAROCFSPSR- 539
DB 479 DHPVISCATKQLRVNGCIPTFTNIGMMVSLRYRKHICGSLIKESVYLTAAROCFSPSR- 538
QY 540 -DLKRYEAMLGIDHVGHGDEKCKQVNLVSQLVYGPESGSDLYLKLARPVLDDFVSTID 598
DB 539 KDLKRYEAMLGIDHVGHGDEKCKQVNLVSQLVYGPESGSDLYLKLARPVLDDFVSTID 598
QY 599 LPNPGCTIPEKTSYVGMGYTGLINYGILRVNHLIYMGNEKCSQHHKGVTLNSESIC 658
DB 599 LPNPGCTIPEKTSYVGMGYTGLINYGILRVNHLIYMGNEKCSQHHKGVTLNSESIC 658
QY 659 AGAKETSGPEEGYGPPLVCEQHKRMVLGVIVPGRGCAIPNRPGLFVAVAYAKMIHK 718
DB 659 AGAKETSGPEEGYGPPLVCEQHKRMVLGVIVPGRGCAIPNRPGLFVAVAYAKMIHK 718
QY 719 IILTYKVPQS 728
DB 719 IILTYKVPQS 728
QY 719 IILTYKVPQS 728
DB 719 IILTYKVPQS 728

OX NCBI_TaxId=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RX MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Bailey S.J., Stern C.D., Gherardi E.;
RT "Expression of HGF/SF, HGF/MSP, and c-met suggests new functions
RL during early chick development.";
RN Dev. Genet. 17:90-101(1995).
RP SEQUENCE OF 1-409 FROM N.A.
RC STRAIN=WHITE LEGHORN, AND RHODE ISLAND RED X LIGHT SUSSEX;
RC TISSUE=EMBRYO;
RX MEDLINE=95237013; PubMed=7720585;
RA Streif A., Stern C.D., Thery C., Ireland G.W., Aparicio S.,
RT "A role for HGF/SF in neural induction and its expression in Hensen's
RL node during gastrulation.";
CC Development 121:813-824(1995).
CC -I SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY
DR EMBL: X84045; CAAS8864.1; -;
DR HSP, X80131; CAAS6430.1; -;
DR HSP, P14210; IBHT.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Trypsin.
DR Pfam: PF00051; Kringle; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_AP; 1.
DR SMART: SM00020; Tryp_Spec; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS0070; KRINGLE_2; 4.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
KW Alternative splicing; Glycoprotein; Growth factor; Hydrolase; Kringle;
FT SIGNAL 1 27
FT CHAIN 28 726
SQ SEQUENCE 726 AA: 82913 MW: 5805F048A576C38 CRC64;
HEPATOCYTE GROWTH FACTOR/ SCATTER FACTOR.
Query Match 76.8%; Score 3170; DB 13; Length 726;
Best Local Similarity 74.5%; Pred. No. 1.4e-280;
Matches 544; Conservative 92; Mismatches 86; Indels 8; Gaps 3;
QY 1 MMTVTLPLALLQVHLHLHLPIAIPYAEGRKRRNTHERKSAKTLIKIDPALKIK 60
DB 1 MMTVTLPLVLLQVHLHLHLHLPIAIPYAEGRKRRNTHERKSAKTLIKIDPALKIK 58
QY 61 TTKVNTADOCANCRNRKNGLPFTCAKAFVDFKARQCLMFPNMSGGKKKEGHEFDLYE 120
DB 59 TTKMTADOCANCRNRKNGLPFTCAKAFVDFKARQCLMFPNMSGGKKKEGHEFDLYE 118
QY 121 NKDYIRNCIIIGKRSYKGTASTITSGIKCOPWSSMTIHEHSFLPSSYRGKDLQENYCRNP 180
DB 119 NKDYIRNCIIIGKRSYKGTASTITSGIKCOPWSSMTIHEHSFLPSSYRGKDLQENYCRNP 178
QY 181 RGEEGPWCFTSNPEVREYVCDIPQCESEVEMTCNGESYRGMDHESGKICQRMHQTP 240
DB 179 RGEEGPWCFTSNPEVREYVCDIPQCESEVEMTCNGESYRGMDHESGKICQRMHQTP 238
QY 241 HRHKEFLPERYPDKGFDDNYCRNPDPGPRWCYTLDPTHEWYCAIKTCADNTMNDTDPVL 300
DB 239 HRHKEFLPERYPDKGFDDNYCRNPDPGPRWCYTLDPTHEWYCAIKTCADNTMNDTDPVL 298
QY 301 ETTECTIOGEGRGYVNTIWNIGIPQCRMDSQYPHEHDMTPENFKCKDLRENYCRNPDS 360
DB 299 ETTECTIOGEGRGYVNTIWNIGIPQCRMDSQYPHEHDMTPENFKCKDLRENYCRNPDS 358

```
Db 297 ETTTCIOGEGYRGVNTWISGICQRMDSOPPHQNHITPENFKCKDLRENYCRNDGS 356
QY 361 ESPWCFTTDDNINRGYSOI.PNCDMSHGQDCYRGNGKNYGNLSQTSGLTCSMDKNME 420
Db 357 ESPWCFTTDDNINRGYSOI.PKCDVSNEDOCYRGNGKSNYGNLSKTRFGLTCSMDKNIE 416
QY 421 DLHRH--IFMEPPASKINENYCRNPDDAHGPMCYTGNPLIPMPCYISRCGGTPTIY 478
Db 417 DLRRHIOIFNEPPVSKKKNYCRNPDDAHGPMCYTGNPLIPMPCYISRCGGTPTIY 476
QY 479 NLDPVISCARKTQLRVNGVCPRTNIGMVSILYRNKHICGSLIKESWVLAAROCFPS 538
Db 477 SLDDTVPICASTKHLRVNVPITQITNGVVSILYRNKHICGSLIKESWVLAAROCFPS 536
QY 539 R--DLKYEAMLGIDHVGHGDECKQVLAWSQLYVPEBSDLVLMKLARAVLDDEVST 596
Db 537 RYDMLKRYKAMLGIDHVGHGDECKQVLAWSQLYVPEBSDLVLMKLARAVLDDEVST 596
QY 597 IDLPNYGCTIPERTSCSVYGMGTGLINVDGLLRVLAHLTYMGNEKCSQHHRGKVTLINESE 656
Db 597 IRLPISGCTIPERTSCSVYGMGTGLINVDGLLRVLAHLTYMGNEKCSQHHRGKVTLINESE 656
QY 657 ICAGAEKIGSGPEGDYGPVLCBOHKRMVLAIVYVPGRCALPNRPGIFVRAVYAKWI 716
Db 657 ICVAETIGAGPCERDYGGLVCEQNRKLIYGVIVPGRCALPNRPGIFVRAVYAKWI 716
QY 717 HKIITFYKVP 726
Db 717 HKIMTYRK 726

RESULT 3
091402 PRELIMINARY; PRT; 710 AA.
AC 091402;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HEPATOCYTE GROWTH FACTOR.
GN HGF.
OS Xenopus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodinae.
OX NCBI_TaxID=8353;
RN [1]
SEQUENCE FROM N.A.
TISUE-TAILED;
MEDLINE=95267690; PubMed=7748783;
Nakamura H., Tashiro K., Nakamura T., Shikawa K.;
"Xenopus cloning of Xenopus HGF cDNA and its expression studies in
RT Molecular cloning of Xenopus HGF cDNA and its expression studies in
RT Mech. Dev. 49:123-131(1995).
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: S77422; AAB34354.2;
DR HSSP: P14210; 1BHT.
DR MEROPS: S01.976;
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Trypsin.
DR Pfam: PF00051; Kringle; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_AP; 1.
DR SMART: SM00020; TRYP_SPE; 1.
DR PROSITE: PS00021; KRINGLE_1; 3.
DR PROSITE: PS50070; KRINGLE_2; 4.
```

```
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 710 AA; 81487 MW; 5FE6480BE31C27FC CRC64;

Query Match 68.0%; Score 2806; DB 13; Length 710;
Best Local Similarity 66.1%; Pred. No. 2,4e-247;
Matches 481; Conservative 107; Mismatches 120; Indels 20; Gaps 5;

QY 1 MWTYKLLPALLQHVLLHLLPLAIPYAEQGRKRRNTIEFKSKATTLIKIDPALKIK 60
Db 1 MWTYKCAWIFEL-----ITLAEGRGKRRNAFDYKKAETTLRLKALEVK 47
QY 61 TKRYNTADOCANRCTRGKLPFTCKAFVDFKAROCIMFPPNSSSGKQKRFGEHPLYE 120
Db 48 TKMFTTENCARKRSRNKGLPFTCKAFVDFKAROCIMFPPNSAGIKDYDISFDLYE 107
QY 121 NKDYIRNCILIGKRSYKGYSTIRKGIKQOPWSSMIPHEHSFLPSYRGKDLQENYCRNP 180
Db 108 KKDVIKRCIHKSGSNYGRTRVTKRGLACQPMNSMIPHEHSFLPSYRGKDLQENYCRNP 167
QY 181 RGEEGPWCFTSNPEVAREVCDIPQSEVBCMTNGESYRGLMDHTESGKICQMDHOTP 240
Db 168 KGEEGPWCFTSNPEVAREVCDIPQSEVBCMTNGESYRGLMDHTESGKICQMDHOTP 227
QY 241 HRHKEFLPERYPDKGFDDNYCRNPDPGPRMPCYTLDPHTRMRYCAIKTCADMTMDTVP 300
Db 228 HRHKEFLPERYPDKGFDDNYCRNPDPGPRMPCYTLDPHTRMRYCAIKTCADMTMDTVP 285
QY 301 ETTTCIOGEGYRGVNTWISGICQRMDSQYVPEHDMPEPFENFKCKDLRENYCRNDGS 360
Db 286 -TKDCMKGGGSGYRGVNTWISGICQRMDSQYVPEHDMPEPFENFKCKDLRENYCRNDGS 344
QY 361 ESPWCFTTDDNINRGYSOI.PNCDMSHGQDCYRGNGKNYGNLSQTSGLTCSMDKNME 420
Db 345 ESPWCFTTDDNINRGYSOI.PKCDVSNEDOCYRGNGKSNYGNLSKTRFGLTCSMDKNIE 404
QY 421 DLHRHIFMEPPASKINENYCRNPDDAHGPMCYTGNPLIPMPCYISRCGGTPTIYNL 480
Db 405 DLKRYEAMLGIDHVGHGDECKQVLAWSQLYVPEBSDLVLMKLARAVLDDEVST 464
QY 481 DHPVISCARKTQLRVNGVCPRTNIGMVSILYRNKHICGSLIKESWVLAAROCFPSR- 539
Db 465 DSP--ITCSSSQQLRVNVPITQITNGVVSILYRNKHICGSLIKESWVLAAROCFPS 523
QY 540 -DLKDYEAMLGIDHVGHGDECKQVLAWSQLYVPEBSDLVLMKLARAVLDDEVST 598
Db 524 IDLYEAMLGIDHVGHGDECKQVLAWSQLYVPEBSDLVLMKLARAVLDDEVST 582
QY 599 LPNYGCTIPERTSCSVYGMGTGLINVDGLLRVLAHLTYMGNEKCSQHHRGKVTLINESE 658
Db 583 LPNYGCTIPERTSCSVYGMGTGLINVDGLLRVLAHLTYMGNEKCSQHHRGKVTLINESE 642
QY 659 AGAEKIGSGPEGDYGPVLCBOHKRMVLAIVYVPGRCALPNRPGIFVRAVYAKWIHK 718
Db 643 AIGETANIGPCERDYGGLVCEQNRKLIYGVIVPGRCALPNRPGIFVRAVYAKWIHK 702
QY 719 IITFYKVP 726
Db 703 IMLTYKAP 710

RESULT 4
091691 PRELIMINARY; PRT; 716 AA.
AC 091691;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE GROWTH FACTOR LIVERINE.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
```

CC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ruiz i Altaba A., Thery C.,
 RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U57455; AAB52574.1; -
 DR HSSP: P00747; ICEA.
 DR MEROPS: S01.00A; -
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR Pfam: PF00051; Kringle.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin.
 DR PRINTS: PR00018; KRINGLE.
 DR SMART: SM00130; KR: 4.
 DR SMART: SM00473; PAN_AP; 1.
 DR SMART: SM00020; tryp_Spc; 1.
 DR PROSITE: PS00021; KRINGLE_1; 4.
 DR PROSITE: PS50070; KRINGLE_2; 4.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 KM Hydrolyase; Serine protease
 SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;

Query Match 41.3%; Score 1704; DB 13; Length 716;
 Best Local Similarity 44.5%; Pred. No. 9, 8e-147;
 Matches 310; Conservative 124; Mismatches 238; Indels 24; Gaps 12;

OY 36 RNTIHEFKSAATTLKIDIPALKIKTKKNTVTDQCANRCKTNGKLPFTCKAFVFPKARKQ 95
 DB 31 RSLNDYQSKGLELVHNNNG-GVKQEIQSEIYQVCAKQSD----LLDRSFVYWMKSGT 85
 OY 96 CLMFPEPNSSGVKKKEFEFLDYENKDIYRNCIIIGKSGSYGTVSTIKSGIKCQPMISM 155
 DB 86 CRLLPWTQNSAVNLLQRRNQYDLYQKKDIYRCVANGNTYIGTSKTSKSGTCCWRMLK 145
 OY 156 IHHHSFLSSYRQKDLQNYCRNRPGEFGPWCTSNPEVRYEVCIDIPOCSEVECMTCN 215
 DB 146 FPHDHKFSPIHW--PELEBNYCRNPDSDPEGPMCTTIDKINHQCGIKKCEDAVCLCN 203
 OY 216 GESYGLMDHTESGKICQICQMDHQTDRHKKFLPERPDGFDNDYCRNPDGQPRPCYILD 275
 DB 204 GEDYGSVDRTSGKECCQRMQDQTPHAPYKPKTPDPSLDNYCRNPDSSERPCYITD 263
 OY 276 PHTREYCAIKTCADNTMNDTVPLETTTCIQQGEGYRGTVNTIWMGIPCQORMDSQYPH 335
 DB 264 PVEKEFCRIYTKCKKQRLSNIEI--TSTCFKREGEYRGKANTTTSGITPCQRNDSQTPQ 320
 OY 336 EHDMPENFKCKDLRENYCRNPDSESPWCTTIDPNTINVCISQIPIKC-DMSHQDCYRG 394
 DB 321 SHRLPEYKPKGIDENYCRNPDSEAPWCFTTLPGRMAVCFQIKRCKDVLBPDCYHG 380
 OY 395 NGKNYMGSLQSTRGLTCSMDKMMEDLHRIHFWEPDASKTENYCRNPDODAGPMPCYT 454
 DB 381 NGELYSGVSTKTRGKICRMRKEKRNDLESLD-QPIYLPLEENYCRNPDSDSGPMCT 439
 OY 455 GNPRLPMDYCIISRCEDGTPTYINLDHPVI--SCAKT----KQLRVANGIPRTNTNGM 508
 DB 440 MDPMPDPDYCAIKPCEKEKYLTLLEASIVYDSCGKRNDRIFFQKSRILVGVF--GNSPWT 497
 OY 509 VSLNVR--NKHICGSLIKESVNLAROCFPS--RDLKDYEMALGIHYVHGDEKCKOYL 565
 DB 498 VSLNROGEHNCGSLVKKENNVISTRCQFSSCDADLSGYEAVMGTLPKNSPDPRDKOSV 557
 OY 566 NVSQLVYEGESDVLNKLKRAVLADVFSTIDLPNGCTIPEKTSQSVYGVGTGLIN 625
 DB 558 PINKIVCGPSSSLVMLKLERPITLNSRVALICLPPRIYVPEETCKEIGKWDTRCTGH 617
 OY 626 DGLRLVNLHYIMGNEKSSQHHRG-KYVLNSETICAGAEKIGSGPCEGDYGGPILVCEQHKM 684

DB 618 DNLKATFIYIISDNCNKNYRSQQRNKVSDNEMCTKPMPIDAGAGEGYGGLACTHDG 677
 OY 685 RAVLGIYVPGCGCAIPNRPDIPFRAVYAKTHKII 720
 DB 678 LVLEGIYVARGCGKKNQPAIFTRVSVYDWINKV 713

RESULT 5
 ID P70006 PRELIMINARY; PRT; 717 AA.
 AC P70006;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN PRECURSOR.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesodactylia; Pipridae; Pipridae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=96404125; PubMed=8808403;
 RA Abergner F., Schmidt G., Richter K.;
 RT "The Xenopus homologue of hepatocyte growth factor-like protein is specifically expressed in the presumptive neural plate during gastrulation."
 RL Mech. Dev. 54:23-37(1996).
 DR EMBL: Y08734; CAA69989.1; -
 DR HSSP: P00747; ICEA.
 DR MEROPS: S01.977; -
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00051; Kringle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00018; KRINGLE.
 DR SMART: SM00130; KR: 4.
 DR SMART: SM00473; PAN_AP; 1.
 DR SMART: SM00020; tryp_Spc; 1.
 DR PROSITE: PS00021; KRINGLE_1; 3.
 DR PROSITE: PS50070; KRINGLE_2; 4.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 KM Hydrolyase; Serine protease; Signal.
 FT SIGNAL 1 28
 FT CHAIN 29 717 HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.
 FT SEQUENCE 717 AA; 82017 MW; 6F877A432C8CD54 CRC64;

Query Match 40.1%; Score 1652.5; DB 13; Length 717;
 Best Local Similarity 42.8%; Pred. No. 4, 9e-142;
 Matches 306; Conservative 126; Mismatches 258; Indels 25; Gaps 13;

OY 17 LHLPLPIAIPYAEQGRKRRNTIHEFKSAATTLKIDIPALKIKTKKNTVTDQCANRCKR 76
 DB 14 LFLCFNAVAFVLTG--TAHRSALNDYQSKGLELVHNNNG-GVKQEVQAEIQCQKQSD 71
 OY 77 NKLPLPTCAFAVFDKARKKOCLEFPENSSGVKKKEFEFLDYENKDIYRNCIIIGKSGY 136
 DB 72 ----LLDCRSFDMNSQSCRLLPMTQNSAVNLLQRRNQYDLYQKKDIYRCVANGNTY 127
 OY 137 KGVSTIKSGICQPMSSNIPHEHSLPSSYRQKDLQNYCRNRPGEFGPWCTSNPEV 196
 DB 128 KGVSTIKSGICQPMSSNIPHEHSLPSSYRQKDLQNYCRNRPGEFGPWCTSNPEV 185
 OY 197 REYVCDIPQCESEVECTNGESYRGLMDHTESGKICQICQMDHQTDRHKKFLPERPDGFD 256
 DB 186 RHQYCGIKKCEDAVCLTNGEDYRGSVDRTSGKECCQRMQDQAPHTHYKPKTPDPSLD 245
 OY 257 DNYCRNPDGQPRPCYITDLPHTREYCAIKTCADNTMNDTVPLETTTCIQQGEGYRG 316

RX MEDLINE-97011126; PubMed-8858136;
RA Onshiro K., Iwama A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,
RA Takasu N., Suda T.;
RT "Molecular cloning of Rat Macrophage-stimulating protein and its
RT involvement in the Male Reproductive System";
RL Biochem. Biophys. Res. Commun. 227:273-280(1996).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR EMBL: X95096; CAA64473.1; -
DR HSSP: P00747; 1KRN.
DR MEROPS: S01.975; -
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan-app.
DR InterPro: IPR001254; Trypsin.
DR Pfam: PF00051; Kringle; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPsin.
DR PRINTS: PR000130; KRINGLE.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_AP; 1.
DR SMART: SM00020; TRYP_SPC; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS50070; KRINGLE_2; 4.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR Hydrolyase; Serine protease; Signal.
KW SIGNA1.
FT SEQUENCE 1 31 POTENTIAL.
SQ 716 AA; 80733 MW; 06B7DF3EF56D921F CRC64;

Query Match 39.1%; Score 1615; DB 11; Length 716;
Best Local Similarity 43.3%; Pred. No. 1.3e-138;

Matches 315; Conservative 118; Mismatches 254; Indels 40; Gaps 16;

QY 17 LHLPLPIAIPYAEGRKRRNTIEFKSAKTLIK-IDPAIKIKTK-VNTADOCANRC 74
DB 4 LPLLLAACSRALGQ---RSPINDFOLRGTELRLNHLHVGVGPMQEDVADAECCARRC 60
QY 75 TRNKGLEPTCKAFVDFKARKKQCLMPFPNSMSSGVKKEFGHEFDLYENKDYIRNCITIGKR 134
DB 61 ---GFLDLCRAFHNYMSSHGCOLLPWTOHSLRAOLHSSLCDLFQKKDYVFTCIDMNGA 116
QY 135 SYKGVTSITKSGIKQCPWMSMIRPHESEFLPSSYRGKDLOENYCRNREGGSGPWCFTSNP 194
DB 117 STRGTARFADGGLPCQAMSRFPNDHKYTPPKNG--LEENFCRNPDGDRGPWCYTTNR 174
QY 195 EVAYEVCDDIPQCESEVEMTCNGESYGLMDHTESGKICQRMWDHQTPIHAKFLPERYPDKG 254
DB 175 SVRFQSCGKIKTCREAVCYLWNGEDYRGEDVYTESGRCQRMWDLOHPSHFPHEKPEPDKA 234
QY 255 FDDNYCRNPDGDRPWCYTLDPHTRMWYCAIKTCADN-----TMDTDVPLETTECIQ 307
DB 235 LKQNYCRNPASERPCWYTTDPVVEREFCDLPSCGRLPPTTKGSKSQQRNKKVASCNFR 294
QY 308 GQEGEYRGVNTIWNGLIPQCRMDSOYRPHEDMTPEMFKCDLRENYCRNPDGSESPWCF 367
DB 295 GKGEDYRGVNTTSAVPCQRMWDQSPHQRVPEKYACKDLRENYCRNPDGSEAPWCF 354
QY 368 TDNINIVGYSQIIPNC-DMSHGQDCYRGNGKNYMGNI,ISFRSLGTCSPMDKXNEDLHRH 426
DB 355 SRGLAVAFQYQIPRCTEEVYRGCGYHSGEDYRGSVSTRKRGVQOQH--SSETPHKPQ 412
QY 427 FWEEDA--SKLNEYCRNPDADAHGWCYTGNDPLIPWYCPISRCGDTPTIVNIDHPV 484
DB 413 FTPTSAFHAGLEAFNFCRNPDGSHGWCYTLDPETLFDYCALRKCDDDPPSI--LDPV 470
QY 485 ----ISGAK----TKOLRVVNGITPIRTNIGMMYSLRR-NKHICGSLINESWLTAROC 535
DB 471 QVOFEKCGKRVNDSNRIRVVGHP--GNSPWTSLRNROGSHFCGSLVEQWVLTAROC 528
QY 536 FPS--RDLKDYFAMLGIHDVHGRGDEKCKOVLNVSOILVYRGESDVLIMKLARPAVLDDF 593

DB 529 IWSCHDPLTGEYEWALGTINQNPQGEANLQRVSAKTVCPAGSQUVLTKLEBPVILNHH 588
QY 594 VSTIDLPNVCCTIPEKTSQSVYGMGTGLINYGDLRVANLYTMGNKCGSHRGKTYLN 633
DB 589 VARICPEQYVPPGPNCEIAGSEKSTNSVTLVHAKMKYISQECWVKRRR--Q 646
QY 654 ESEICAGAEKIGSGPCGDDGGLPVCQEHKMRVLYGVTRGCAIPNRPGIFRVAYYA 713
DB 647 ESEICTGGLAPTAGCAGDYGGPLACTTHDCWVLOGILINRVACARPRPAITRVSVEV 706
QY 714 KWIHKII 720
DB 707 DWINKV 713
RESULT 10
Q91XG8 PRELIMINARY; PRT; 716 AA.
ID Q91XG8
AC Q91XG8;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE HEPATOCYTE GROWTH FACTOR-LIKE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Strausberg R.;
RU Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC010551; AAL10551.1; -
SQ SEQUENCE 716 AA; 80693 MW; 12474C48A7D4B46D CRC64;

Query Match 38.9%; Score 1605; DB 11; Length 716;
Best Local Similarity 43.0%; Pred. No. 1.1e-137;

Matches 312; Conservative 113; Mismatches 263; Indels 38; Gaps 14;

QY 17 LHLPLPIAIPYAEGRKRRNTIEFKSAKTLIKI--DPAIKIKTKVNTADOCANR 73
DB 4 LPLLLLVQCSRALGQSRPLNDFOFLRGTELRLNHLHVAVGP---WQEDVADAECCARR 59
QY 74 CTRNKGLEPTCKAFVDFKARKKQCLMPFPNSMSSGVKKEFGHEFDLYENKDYIRNCITIGK 133
DB 60 C---GFLDLCRAFHNYMSSHGCOLLPWTOHSLTQLYHSSLCDLFQKKDYVFTCIDMNG 115
QY 134 RSYKGVTSITKSGIKQCPWMSMIRPHESEFLPSSYRGKDLOENYCRNREGGSGPWCFTSNP 193
DB 116 VSTRGTARFADGGLPCQAMSRFPNDHKYTPPKNG--LEENFCRNPDGDRGPWCYTTNR 173
QY 194 PEVAYEVCDDIPQCESEVEMTCNGESYGLMDHTESGKICQRMWDHQTPIHAKFLPERYPDK 253
DB 174 RSVRFQSCGKIKTCREAVCYLWNGEDYRGEDVYTESGRCQRMWDLOHPSHFPHEKPEPDK 233
QY 254 GPDNYCRNPDGDRPWCYTLDPHTRMWYCAIKTCADN-----TMDTDVPLETTECI 306
DB 234 DLKQNYCRNPDGSESPWYTTDPVVEREFCDLPSCGRLPPTTKGSKSQQRNKKVASCNFR 293
QY 307 QOGEYRGVNTIWNGLIPQCRMDSOYRPHEDMTPEMFKCDLRENYCRNPDGSESPWCF 366
DB 294 GKGEDYRGVNTTSAVPCQRMWDQSPHQRVPEKYACKDLRENYCRNPDGSEAPWCF 353
QY 367 TTDNINIVGYSQIIPNC-DMSHGQDCYRGNGKNYMGNI,ISFRSLGTCSPMDKXNEDLHRH 425
DB 354 TSRLGLMAFCHOQIPRCTEEVYRGCGYHSGEDYRGSVSTRKRGVQOQH--SSETPHKP 411
QY 426 IWEEDASK--LNEYCRNPDADAHGWCYTGNDPLIPWYCPISRCGDTPTIVNIDHP 483
DB 412 QFTPTSAFHAGLEAFNFCRNPDGSHGWCYTLDPETLFDYCALRKCDDDPPSI--LDPV 471

Qy	484	VI	SCA-----KTQLRVNVIPIPTRTINGMNVSLRYR-NKHICGSGSLIKESVWLTAAOCF	536
Db	472	VFEKCGKRVYDKSNKLRVVGSHR--GNSPWTVSLRNNGQHFCCGSLVYKQWVLTAAOC	529	
Qy	537	PS--RDLEKDYEAWLGIHDVHGRGDEKCKQVLTNSQLVYVPEGSDDLMLKARPAVLDEY	594	
Db	530	MSCHERPLTGVLEVWLTGINQNPORBANQVRVAKVCGRPGSQLVTLKIERVLTNNHY	569	
Qy	555	STIDLPNVCSTIPREKTSVYGVKGTGLINTDGLLRVNAHLITMGNEKCSOHNRKGTVLNE	654	
Db	550	ALICLPBPQVYVPRQTKELGWMGEESTIGTSNNVTLHNASMNVISNOCENKRYRGHT--QE	647	
Qy	655	SEICGAEEKIGSGPECGDYGGPLVCEQHKMMVLGVITVPGCAIPNRPGRIVVAYYAK	714	
Db	648	SEICQGMNVVYVACESBGYGGPLACTYHDCNVLDGGLIPNRVCARPRMPALFRVFRVYVD	707	
Qy	715	WIKTTI	720	
Db	708	WINKVM	713	

DT	ID	Q902N6	PRELIMINARY;	PRT;	709 AA.
AC	Q902N6;				
DT	01-DEC-2001	(TREMBLrel. 19, Created)			
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			
DE	HEPATOCYTE GROWTH FACTOR-LIKE 1.				
GN	HGF1L.				
OS	Brachydanio rerio (zebrafish) (zebra danio).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;				
OC	Cypriniformes; Cyprinidae; Danio.				
OX	NCBI_TaxID=7955;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Bassett D.I., Wilson S.W.;				
RT	"Early expression of zebrafish Hepatocyte Growth Factor-Like 1				
RT	suggests a conserved role in vertebrate neural induction."				
RL	Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF370035; AAK54207.1; -				
SO	SEQUENCE 709 AA; 81271 MW; 99072365D873A20 CRC64;				

Query Match	38.3%;	Score 1579;	DB 13;	Length 709;
Best Local Similarity	42.08;	Pred. No. 2.5e-135;		
Matches 312;	Conservative 110;	Mismatches 245;	Indels 76;	Gaps 19;

[illegible][illegible]

RESULT	12		
ID	018783	PRELIMINARY;	PRT; 806 AA.
AC	018783;		
DT	01-JAN-1998 (TrEMBLrel. 05, Created)		
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)		
DE	PLASMINOGEN.		
OS	Macropus eugenii (Tamar wallaby).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.		
NC	NCBI_TaxId=9315;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=LIVER.		
RX	MEDLINE=98004511; PubMed=9342350;		
RA	Lawn R.M., Schwartz K., Patthy L.;		
RL	"Convergent evolution of apolipoprotein(a) in primates and hedgehog."		
RT	Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE		
CC	TRYPSIN FAMILY.		
DR	EMBL; AF012297; AAB65760.1; -.		
DR	HSSP; P00747; SHPG.		
DR	MEROPS; S01_233; -.		
DR	InterPro; IPR003134; Chymotrypsin.		
DR	InterPro; IPR000001; Kringle.		
DR	InterPro; IPR0035014; PAN.		
DR	InterPro; IPR003609; Pan_app.		
DR	InterPro; IPR001254; Trypsin.		
DR	Pfam; PF00051; kringle; 5.		
DR	Pfam; PF00024; PAN; 1.		
DR	Pfam; PF00089; trypsin; 1.		
DR	PRINTS; PR00722; CHYMOTRYPsin.		
DR	PRINTS; PR00018; KRINGLE.		
DR	SMART; SM00130; KR; 4.		
DR	SMART; SM00473; PAN_AP; 1.		
DR	SMART; SM00020; TRYp_SPC; 1.		
DR	PROSITE; PS00021; KRINGLE.1; 5.		
DR	PROSITE; PSS0070; KRINGLE.2; 5.		
DR	PROSITE; PSS0240; TRYPSIN_DOM; 1.		
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.		
DR	PROSITE; PS00135; TRYPSIN_SER; 1.		
DR	Hydrolase; Serine protease.		
SW	SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;		


```

QY 61 TRKVNADOCANCRTRNKGLEPTCAKAFVDFKARKOCLMEFPNMSGGVKKKEGHEFDLYE 120
DB 61 TRKVNADOCANCRTRNKGLEPTCAKAFVDFKARKOCLMEFPNMSGGVKKKEGHEFDLYE 120
QY 121 NKDYINCLIGGRSYKGVSTITSKIKQCPMSSMIPHEHSFLPSSYRGKDLQENYCRNP 180
DB 121 NKDYINCLIGGRSYKGVSTITSKIKQCPMSSMIPHEHSFLPSSYRGKDLQENYCRNP 180
QY 181 RKEEGGPMCFSTNPEVRYEVCIDPQSEVECMTCNGESYRGMLDHTESKICQRMWHQTP 240
DB 181 RKEEGGPMCFSTNPEVRYEVCIDPQSEVECMTCNGESYRGMLDHTESKICQRMWHQTP 240
QY 241 HRRKFLPERYPDKGFDNNYCRNPDDAGHPWCYTGNPLIPMDYCPISRCGDDTPTIVNL 300
DB 241 HRRKFLPERYPDKGFDNNYCRNPDDAGHPWCYTGNPLIPMDYCPISRCGDDTPTIVNL 300
QY 301 ETTECIQGGEGYRGVNTINMGICQQRWDSQYPHEHDMTPENFKCKDLRENYCRNPDS 360
DB 301 ETTECIQGGEGYRGVNTINMGICQQRWDSQYPHEHDMTPENFKCKDLRENYCRNPDS 360
QY 361 ESPMCFETTDPIRNVGYCSQIPNCDMSHGDDCYRGNGKNYMGNLISOTRSGLTCSMDXNME 420
DB 361 ESPMCFETTDPIRNVGYCSQIPNCDMSHGDDCYRGNGKNYMGNLISOTRSGLTCSMDXNME 420
QY 421 DLHRRHIFWEPDASKLNNYCRNPDDAGHPWCYTGNPLIPMDYCPISRCGDDTPTIVNL 480
DB 421 DLHRRHIFWEPDASKLNNYCRNPDDAGHPWCYTGNPLIPMDYCPISRCGDDTPTIVNL 480
QY 481 DHPVISCATKQRLRVVNGIPRTNIGMWSLRYRNKHICGSLIKESVWLTARQCFPSRD 540
DB 481 DHPVISCATKQRLRVVNGIPRTNIGMWSLRYRNKHICGSLIKESVWLTARQCFPSRD 540
QY 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDYLMLKARPAVLDDFVSTIDLP 600
DB 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDYLMLKARPAVLDDFVSTIDLP 600
QY 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYMGNEKCSQHHRKQVLTLESEICAG 660
DB 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYMGNEKCSQHHRKQVLTLESEICAG 660
QY 661 AEKISGPGCEGDYGGPLVCEQHKKRMVGLVIVPGRCALIPNRGIFVRAVYAKWIIKII 720
DB 661 AEKISGPGCEGDYGGPLVCEQHKKRMVGLVIVPGRCALIPNRGIFVRAVYAKWIIKII 720
QY 721 LTYKVPQS 728
DB 721 LTYKVPQS 728

```

```

PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/884811
FILING DATE: 18-MAY-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/885971
FILING DATE: 18-MAY-92
ATTORNEY/AGENT INFORMATION:
NAME: Marschang, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P0755779P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5416
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 728 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-087-783A-22

Query Match      100.0%; Score 4126; DB 1; Length 728;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMYVTLRALLQHVHLHLPLIPVYAEGRKRRNTIHEFKSAKTTLIKIDPALKIK 60
DB 1 MMYVTLRALLQHVHLHLPLIPVYAEGRKRRNTIHEFKSAKTTLIKIDPALKIK 60
QY 61 TRKVNADOCANCRTRNKGLEPTCAKAFVDFKARKOCLMEFPNMSGGVKKKEGHEFDLYE 120
DB 61 TRKVNADOCANCRTRNKGLEPTCAKAFVDFKARKOCLMEFPNMSGGVKKKEGHEFDLYE 120
QY 121 NKDYINCLIGGRSYKGVSTITSKIKQCPMSSMIPHEHSFLPSSYRGKDLQENYCRNP 180
DB 121 NKDYINCLIGGRSYKGVSTITSKIKQCPMSSMIPHEHSFLPSSYRGKDLQENYCRNP 180
QY 181 RKEEGGPMCFSTNPEVRYEVCIDPQSEVECMTCNGESYRGMLDHTESKICQRMWHQTP 240
DB 181 RKEEGGPMCFSTNPEVRYEVCIDPQSEVECMTCNGESYRGMLDHTESKICQRMWHQTP 240
QY 241 HRRKFLPERYPDKGFDNNYCRNPDDAGHPWCYTGNPLIPMDYCPISRCGDDTPTIVNL 300
DB 241 HRRKFLPERYPDKGFDNNYCRNPDDAGHPWCYTGNPLIPMDYCPISRCGDDTPTIVNL 300
QY 301 ETTECIQGGEGYRGVNTINMGICQQRWDSQYPHEHDMTPENFKCKDLRENYCRNPDS 360
DB 301 ETTECIQGGEGYRGVNTINMGICQQRWDSQYPHEHDMTPENFKCKDLRENYCRNPDS 360
QY 361 ESPMCFETTDPIRNVGYCSQIPNCDMSHGDDCYRGNGKNYMGNLISOTRSGLTCSMDXNME 420
DB 361 ESPMCFETTDPIRNVGYCSQIPNCDMSHGDDCYRGNGKNYMGNLISOTRSGLTCSMDXNME 420
QY 421 DLHRRHIFWEPDASKLNNYCRNPDDAGHPWCYTGNPLIPMDYCPISRCGDDTPTIVNL 480
DB 421 DLHRRHIFWEPDASKLNNYCRNPDDAGHPWCYTGNPLIPMDYCPISRCGDDTPTIVNL 480
QY 481 DHPVISCATKQRLRVVNGIPRTNIGMWSLRYRNKHICGSLIKESVWLTARQCFPSRD 540
DB 481 DHPVISCATKQRLRVVNGIPRTNIGMWSLRYRNKHICGSLIKESVWLTARQCFPSRD 540
QY 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDYLMLKARPAVLDDFVSTIDLP 600
DB 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDYLMLKARPAVLDDFVSTIDLP 600
QY 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYMGNEKCSQHHRKQVLTLESEICAG 660
DB 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYMGNEKCSQHHRKQVLTLESEICAG 660
QY 661 AEKISGPGCEGDYGGPLVCEQHKKRMVGLVIVPGRCALIPNRGIFVRAVYAKWIIKII 720
DB 661 AEKISGPGCEGDYGGPLVCEQHKKRMVGLVIVPGRCALIPNRGIFVRAVYAKWIIKII 720

```


APPLICANT: TSUDA, EISUKE
APPLICANT: MASUNAGA, HIRAKI
APPLICANT: TAKAHARA, REIKO
APPLICANT: OOGAKI, FUMIKO
APPLICANT: UEDA, MASATSUGU
APPLICANT: HIGASHIO, KANJI
TITLE OF INVENTION: MODIFIED TCF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault
STREET: 125 High St.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,937B
FILING DATE: 19-AUG-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A.
REGISTRATION NUMBER: 32,503
REFERENCE/DOCKET NUMBER: FJN-022
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-937b-1

Query Match 99.1%; Score 4088.5; DB 1; Length 723;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

1 MMVTKRLPALLOVHLHLPLTAIPAEQORRRNTIHEFKSAKTLKIDPALKIK 60
1 MMTVKRLPALLOVHLHLPLTAIPAEQORRRNTIHEFKSAKTLKIDPALKIK 60
61 TKKVTADOCANRCTRNKGLPFTCKAFVFDKARKOCLMFPNSMSSGVKKFEGHEFDLYE 120
61 TKKVTADOCANRCTRNKGLPFTCKAFVFDKARKOCLMFPNSMSSGVKKFEGHEFDLYE 120
61 TKKVTADOCANRCTRNKGLPFTCKAFVFDKARKOCLMFPNSMSSGVKKFEGHEFDLYE 120
121 NKDYIRNCIIGKRSYKTSVITSGIKQCPWSSNIPIHESFLPSSYKGLQENYCNRP 180
121 NKDYIRNCIIGKRSYKTSVITSGIKQCPWSSNIPIHESFLPSSYKGLQENYCNRP 180
121 NKDYIRNCIIGKRSYKTSVITSGIKQCPWSSNIPIHESFLPSSYKGLQENYCNRP 180
181 RGEESGPMCFSTNPEVREVDIPQCSVEECMTGSEYRGIMHTESGKICQMRDHTP 240
181 RGEESGPMCFSTNPEVREVDIPQCSVEECMTGSEYRGIMHTESGKICQMRDHTP 240
176 RGEESGPMCFSTNPEVREVDIPQCSVEECMTGSEYRGIMHTESGKICQMRDHTP 235
241 HRHFLPERYPDKGFDNDYCNRPDQPRPWCYTLDPRHMEYCAIKTCADNTMNDTVPL 300
241 HRHFLPERYPDKGFDNDYCNRPDQPRPWCYTLDPRHMEYCAIKTCADNTMNDTVPL 300
236 HRHFLPERYPDKGFDNDYCNRPDQPRPWCYTLDPRHMEYCAIKTCADNTMNDTVPL 295
301 ETTKICIGOGGEGYGTVNTINNGIIPCQWRMSQYRPHEDMTPEENFKCKDLRENYCNPDS 360
301 ETTKICIGOGGEGYGTVNTINNGIIPCQWRMSQYRPHEDMTPEENFKCKDLRENYCNPDS 360
296 ETTKICIGOGGEGYGTVNTINNGIIPCQWRMSQYRPHEDMTPEENFKCKDLRENYCNPDS 355
361 ESPPCFTDPNIRVGYGQINCMDSHGODCYRGNGKRYMGNLSOTRSGLTCSMDKME 420
361 ESPPCFTDPNIRVGYGQINCMDSHGODCYRGNGKRYMGNLSOTRSGLTCSMDKME 420
356 ESPPCFTDPNIRVGYGQINCMDSHGODCYRGNGKRYMGNLSOTRSGLTCSMDKME 415

421 DLHRIHWEPPASKLENYCNRPDDAHGPMCTGNPLIPMDYCPISRCEDDTPTIYNL 480
416 DLHRIHWEPPASKLENYCNRPDDAHGPMCTGNPLIPMDYCPISRCEDDTPTIYNL 475
481 DHPVISCARTKOLRVANGIPTRTNIGMWSLRYRNKHIGGSLIKESWVLTAROCFSPSRD 540
476 DHPVISCARTKOLRVANGIPTRTNIGMWSLRYRNKHIGGSLIKESWVLTAROCFSPSRD 535
541 LKDYKAMGIDHVGREGCKOVYLVNSQVLYGPGSDLYIMKLARPAVLDDFVSTIDLP 600
536 LKDYKAMGIDHVGREGCKOVYLVNSQVLYGPGSDLYIMKLARPAVLDDFVSTIDLP 595
601 NYGCTIPEKTSQSVYGMGYTGLINVDGLRYVAHLIYMGNEKCSQHRGKVTLNSEICAG 660
596 NYGCTIPEKTSQSVYGMGYTGLINVDGLRYVAHLIYMGNEKCSQHRGKVTLNSEICAG 655
661 AEKIGSGPCEDYGGPLVCEQHKMRVLCVYVPGRCALPVRPGIFVAVAYAKIRKII 720
656 AEKIGSGPCEDYGGPLVCEQHKMRVLCVYVPGRCALPVRPGIFVAVAYAKIRKII 715
721 LTKVYPOS 728
716 LTKVYPOS 723

RESULT 5

US-08-404-643-1
Sequence 1, Application US/08404643
Patent No. 5658742
GENERAL INFORMATION:
APPLICANT: HIGASHIO, KANJI
APPLICANT: SHIMA, NOBUYUKI
APPLICANT: OOGAKI, FUMIKO
TITLE OF INVENTION: MONOCLONAL ANTIBODY
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
STREET: 53 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/404,643
FILING DATE: 15-MAR-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A.
REGISTRATION NUMBER: 32,503
REFERENCE/DOCKET NUMBER: FJN-034 (3999/35)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)-248-7000
TELEFAX: (617)-248-7100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-404-643-1

Query Match 99.1%; Score 4088.5; DB 1; Length 723;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 MMTVTLRALLQVHLHLPLPAIPYAEQRRRNTIHEFFKSATTLIKIDPALKIK 60
; SEQUENCE ID NO. 1
; LENGTH: 723
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: TCF-II
US-09-194-326-1

Db 1 MMTVTLRALLQVHLHLPLPAIPYAEQRRRNTIHEFFKSATTLIKIDPALKIK 60
1 MMTVTLRALLQVHLHLPLPAIPYAEQRRRNTIHEFFKSATTLIKIDPALKIK 60

QY 61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPNSMSGYKKEGHEFDLYE 120
61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPNSMSGYKKEGHEFDLYE 120

Db 61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPNSMSGYKKEGHEFDLYE 120
61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPNSMSGYKKEGHEFDLYE 120

QY 121 NKDYIRNCIIIGKGRSYKTVSITKSGIKQCPWSSMIPEHSFLSSYRGKDLQENYCRNP 180
121 NKDYIRNCIIIGKGRSYKTVSITKSGIKQCPWSSMIPEHSFLSSYRGKDLQENYCRNP 180

Db 121 NKDYIRNCIIIGKGRSYKTVSITKSGIKQCPWSSMIPEHSFLSSYRGKDLQENYCRNP 180
121 NKDYIRNCIIIGKGRSYKTVSITKSGIKQCPWSSMIPEHSFLSSYRGKDLQENYCRNP 180

QY 181 RGEEGGPMCTSNPEVREYECDDIPQCESEVEMTCNGESYRGIMDHTEGSKICQRMHDOTP 240
181 RGEEGGPMCTSNPEVREYECDDIPQCESEVEMTCNGESYRGIMDHTEGSKICQRMHDOTP 240

Db 181 RGEEGGPMCTSNPEVREYECDDIPQCESEVEMTCNGESYRGIMDHTEGSKICQRMHDOTP 240
181 RGEEGGPMCTSNPEVREYECDDIPQCESEVEMTCNGESYRGIMDHTEGSKICQRMHDOTP 240

QY 241 HRHFLPERYPDKGFDDNYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 300
241 HRHFLPERYPDKGFDDNYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 300

Db 241 HRHFLPERYPDKGFDDNYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 300
241 HRHFLPERYPDKGFDDNYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 300

QY 301 ETTECIGOGGEGYRGTVNTIINGIIPCQRMDSQYRHEHDMTEPNFKCDLRNRYCRNPDGS 360
301 ETTECIGOGGEGYRGTVNTIINGIIPCQRMDSQYRHEHDMTEPNFKCDLRNRYCRNPDGS 360

Db 301 ETTECIGOGGEGYRGTVNTIINGIIPCQRMDSQYRHEHDMTEPNFKCDLRNRYCRNPDGS 360
301 ETTECIGOGGEGYRGTVNTIINGIIPCQRMDSQYRHEHDMTEPNFKCDLRNRYCRNPDGS 360

QY 361 ESPMCFETDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNLSTQTRSGLTCSMMDKME 420
361 ESPMCFETDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNLSTQTRSGLTCSMMDKME 420

Db 361 ESPMCFETDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNLSTQTRSGLTCSMMDKME 420
361 ESPMCFETDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNLSTQTRSGLTCSMMDKME 420

QY 421 DLHRIEWPDPASKLNEYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 480
421 DLHRIEWPDPASKLNEYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 480

Db 421 DLHRIEWPDPASKLNEYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 480
421 DLHRIEWPDPASKLNEYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 480

QY 481 DHPVISCARTKQLRVVNGIPIRTNIGMWVSLRYRNKHIICGSSLIKESVNLAROCFSPSRD 540
481 DHPVISCARTKQLRVVNGIPIRTNIGMWVSLRYRNKHIICGSSLIKESVNLAROCFSPSRD 540

Db 481 DHPVISCARTKQLRVVNGIPIRTNIGMWVSLRYRNKHIICGSSLIKESVNLAROCFSPSRD 540
481 DHPVISCARTKQLRVVNGIPIRTNIGMWVSLRYRNKHIICGSSLIKESVNLAROCFSPSRD 540

QY 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600
541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600

Db 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600
541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600

QY 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSQHHKRGVTLNSEICAG 660
601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSQHHKRGVTLNSEICAG 660

Db 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSQHHKRGVTLNSEICAG 660
601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSQHHKRGVTLNSEICAG 660

QY 661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGGCAIPRRPGIFVAVAYAAWIHKII 720
661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGGCAIPRRPGIFVAVAYAAWIHKII 720

Db 661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGGCAIPRRPGIFVAVAYAAWIHKII 720
661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGGCAIPRRPGIFVAVAYAAWIHKII 720

QY 721 LTYKVPQS 728
721 LTYKVPQS 728

Db 721 LTYKVPQS 728
721 LTYKVPQS 728

QY 716 LTYKVPQS 723
716 LTYKVPQS 723

Db 716 LTYKVPQS 723
716 LTYKVPQS 723

RESULT 6
US-09-194-326-1
; Sequence 1, Application US/09194326
; Patent No. 6306827
; GENERAL INFORMATION:
; APPLICANT: Kinoshita, Masahiko
; APPLICANT: Ogawa, Hiroshi
; APPLICANT: Masanaga, Hiroaki
; APPLICANT: Kobayashi, Fumie
; APPLICANT: Yamaguchi, Kyoji
; APPLICANT: Higashio, Kanji
; TITLE OF INVENTION: Agent for Preventing and/or Treating Renal Disease
; FILE REFERENCE: FPN-069
; CURRENT APPLICATION NUMBER: US/09/194,326
; EARLIER APPLICATION NUMBER: PCT/JP98/01221
; EARLIER FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: JP 94989
; EARLIER FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 3

QY 1 MMTVTLRALLQVHLHLPLPAIPYAEQRRRNTIHEFFKSATTLIKIDPALKIK 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO. 1
; LENGTH: 723
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: TCF-II
US-09-194-326-1

Db 1 MMTVTLRALLQVHLHLPLPAIPYAEQRRRNTIHEFFKSATTLIKIDPALKIK 60
1 MMTVTLRALLQVHLHLPLPAIPYAEQRRRNTIHEFFKSATTLIKIDPALKIK 60

QY 61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPNSMSGYKKEGHEFDLYE 120
61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPNSMSGYKKEGHEFDLYE 120

Db 61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPNSMSGYKKEGHEFDLYE 120
61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPNSMSGYKKEGHEFDLYE 120

QY 121 NKDYIRNCIIIGKGRSYKTVSITKSGIKQCPWSSMIPEHSFLSSYRGKDLQENYCRNP 180
121 NKDYIRNCIIIGKGRSYKTVSITKSGIKQCPWSSMIPEHSFLSSYRGKDLQENYCRNP 180

Db 121 NKDYIRNCIIIGKGRSYKTVSITKSGIKQCPWSSMIPEHSFLSSYRGKDLQENYCRNP 180
121 NKDYIRNCIIIGKGRSYKTVSITKSGIKQCPWSSMIPEHSFLSSYRGKDLQENYCRNP 180

QY 181 RGEEGGPMCTSNPEVREYECDDIPQCESEVEMTCNGESYRGIMDHTEGSKICQRMHDOTP 240
181 RGEEGGPMCTSNPEVREYECDDIPQCESEVEMTCNGESYRGIMDHTEGSKICQRMHDOTP 240

Db 181 RGEEGGPMCTSNPEVREYECDDIPQCESEVEMTCNGESYRGIMDHTEGSKICQRMHDOTP 240
181 RGEEGGPMCTSNPEVREYECDDIPQCESEVEMTCNGESYRGIMDHTEGSKICQRMHDOTP 240

QY 241 HRHFLPERYPDKGFDDNYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 300
241 HRHFLPERYPDKGFDDNYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 300

Db 241 HRHFLPERYPDKGFDDNYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 300
241 HRHFLPERYPDKGFDDNYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 300

QY 301 ETTECIGOGGEGYRGTVNTIINGIIPCQRMDSQYRHEHDMTEPNFKCDLRNRYCRNPDGS 360
301 ETTECIGOGGEGYRGTVNTIINGIIPCQRMDSQYRHEHDMTEPNFKCDLRNRYCRNPDGS 360

Db 301 ETTECIGOGGEGYRGTVNTIINGIIPCQRMDSQYRHEHDMTEPNFKCDLRNRYCRNPDGS 360
301 ETTECIGOGGEGYRGTVNTIINGIIPCQRMDSQYRHEHDMTEPNFKCDLRNRYCRNPDGS 360

QY 361 ESPMCFETDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNLSTQTRSGLTCSMMDKME 420
361 ESPMCFETDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNLSTQTRSGLTCSMMDKME 420

Db 361 ESPMCFETDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNLSTQTRSGLTCSMMDKME 420
361 ESPMCFETDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNLSTQTRSGLTCSMMDKME 420

QY 421 DLHRIEWPDPASKLNEYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 480
421 DLHRIEWPDPASKLNEYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 480

Db 421 DLHRIEWPDPASKLNEYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 480
421 DLHRIEWPDPASKLNEYCRNPDGAPRPMWCYTLDPRHMEYCAIKTCADNTMDTVPL 480

QY 481 DHPVISCARTKQLRVVNGIPIRTNIGMWVSLRYRNKHIICGSSLIKESVNLAROCFSPSRD 540
481 DHPVISCARTKQLRVVNGIPIRTNIGMWVSLRYRNKHIICGSSLIKESVNLAROCFSPSRD 540

Db 481 DHPVISCARTKQLRVVNGIPIRTNIGMWVSLRYRNKHIICGSSLIKESVNLAROCFSPSRD 540
481 DHPVISCARTKQLRVVNGIPIRTNIGMWVSLRYRNKHIICGSSLIKESVNLAROCFSPSRD 540

QY 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600
541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600

Db 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600
541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600

QY 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSQHHKRGVTLNSEICAG 660
601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSQHHKRGVTLNSEICAG 660

Db 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSQHHKRGVTLNSEICAG 660
601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSQHHKRGVTLNSEICAG 660

QY 661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGGCAIPRRPGIFVAVAYAAWIHKII 720
661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGGCAIPRRPGIFVAVAYAAWIHKII 720

Db 661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGGCAIPRRPGIFVAVAYAAWIHKII 720
661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGGCAIPRRPGIFVAVAYAAWIHKII 720

QY 721 LTYKVPQS 728
721 LTYKVPQS 728

Db 721 LTYKVPQS 728
721 LTYKVPQS 728

QY 716 LTYKVPQS 723
716 LTYKVPQS 723

Db 716 LTYKVPQS 723
716 LTYKVPQS 723

RESULT 7
US-08-290-937B-2
; Sequence 2, Application US/08290937B
; Patent No. 5648233
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, KYOJI

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-937B-3

Query Match 98.9%; Score 4079.5; DB 1; Length 723;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 721; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

```

QY 1 MWYTKLIPALLQVHLHLHLPLAIPYAGQRRKRNTHIEFKSATTTLIKIDPALKIK 60
   |||||||
DB 1 MWYTKLIPALLQVHLHLHLPLAIPYAGQRRKRNTHIEFKSATTTLIKIDPALKIK 60
QY 61 TKKYNADQCANRCTRNKGLPFTCKAFVFDKARKQCLMFPPNSMSGYKKEFGHEFDLYE 120
   |||||||
DB 61 TKKYNADQCANRCTRNKGLPFTCKAFVFDKARKQCLMFPPNSMSGYKKEFGHEFDLYE 120
QY 121 NKDIYRNCIIIGKGRSYKGVSTIKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYCRNP 180
   |||||||
DB 121 NKDIYRNCIIIGKGRSYKGVSTIKSGIKQCPWSSMIPHEH-----SYRGKDLQENYCRNP 175
QY 181 RGEGBGPMCTSNPEVREYECDDIPQSEVECMTCNGESYRGKDLMDHTSSGKICQRMWDQTP 240
   |||||||
DB 176 RGEGBGPMCTSNPEVREYECDDIPQSEVECMTCNGESYRGKDLMDHTSSGKICQRMWDQTP 235
QY 241 HRHKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDPTRMWEYCAIKTCADNTMDTDVPL 300
   |||||||
DB 236 HRHKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDPTRMWEYCAIKTCADNTMDTDVPL 295
QY 301 ETTECIOGQGBGYRGVNTWNGIIPCQRMDSQYRHEHDMTPENFKCKDLRENYCRNPDGS 360
   |||||||
DB 296 ETTECIOGQGBGYRGVNTWNGIIPCQRMDSQYRHEHDMTPENFKCKDLRENYCRNPDGS 355
QY 361 ESPMCFETTDNIRVGYCSQIPNCDMSHGQDCYRGNGKNYGNLSQTSRGLTCSMWDKME 420
   |||||||
DB 356 ESPMCFETTDNIRVGYCSQIPNCDMSHGQDCYRGNGKNYGNLSQTSRGLTCSMWDKME 415
QY 421 DLHRHIFWEPDASKLNEYCRNPDDDAHGPMCYTGNPLIPDYCPISRCBEDTPTIYNL 480
   |||||||
DB 416 DLHRHIFWEPDASKLNEYCRNPDDDAHGPMCYTGNPLIPDYCPISRCBEDTPTIYNL 475
QY 481 DHPVISCARKTKOLRVNGIPLRTNIGMMVSLRYRNKHIICGSLIKESMWLTARQCFPSRD 540
   |||||||
DB 476 DHPVISCARKTKOLRVNGIPLRTNIGMMVSLRYRNKHIICGSLIKESMWLTARQCFPSRD 535
QY 541 LKQYEANLGIHDVHGRDECKQVLYNSQVYGGSDVLYMLKARAVLDDEFSTIDLP 600
   |||||||
DB 536 LKQYEANLGIHDVHGRDECKQVLYNSQVYGGSDVLYMLKARAVLDDEFSTIDLP 595
QY 601 NYGCTIPEKTSYVGMGTGLINYDGLRVYAHLYIMGNEKCSOHHHGKVTLLNESEICAG 660
   |||||||
DB 596 NYGCTIPEKTSYVGMGTGLINYDGLRVYAHLYIMGNEKCSOHHHGKVTLLNESEICAG 655
QY 661 AEKIGSPCEGDYGGPLVCEQHKMRYLVGIIVPGGCAIPNRPGLIFRVAVYAKMIHKII 720
   |||||||
DB 656 AEKIGSPCEGDYGGPLVCEQHKMRYLVGIIVPGGCAIPNRPGLIFRVAVYAKMIHKII 715
QY 721 LTYKVPQS 728
   |||||||
DB 716 LTYKVPQS 723

```

RESULT 9
US-08-030-410-3
Sequence 3, Application US/08030410
Patent No. 6221359

GENERAL INFORMATION:
APPLICANT: Komiyama, Atsushi
APPLICANT: Nakahata, Tatsutoshi
APPLICANT: Kubo, Tetsuo
APPLICANT: Tanaka, Ryunei
APPLICANT: Kawano, Genji
APPLICANT: Sudo, Tetsuo

```

APPLICANT: Sano, Emiko
APPLICANT: Kojima, Katsunaki
TITLE OF INVENTION: HEMOPOIETIC STEM CELL MULTIPLIER
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/030,410
FILING DATE: 19930521
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8898
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-030-410-3

```

Query Match 98.8%; Score 4077.5; DB 4; Length 723;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 720; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

```

QY 1 MWYTKLIPALLQVHLHLHLPLAIPYAGQRRKRNTHIEFKSATTTLIKIDPALKIK 60
   |||||||
DB 1 MWYTKLIPALLQVHLHLHLPLAIPYAGQRRKRNTHIEFKSATTTLIKIDPALKIK 60
QY 61 TKKYNADQCANRCTRNKGLPFTCKAFVFDKARKQCLMFPPNSMSGYKKEFGHEFDLYE 120
   |||||||
DB 61 TKKYNADQCANRCTRNKGLPFTCKAFVFDKARKQCLMFPPNSMSGYKKEFGHEFDLYE 120
QY 121 NKDIYRNCIIIGKGRSYKGVSTIKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYCRNP 180
   |||||||
DB 121 NKDIYRNCIIIGKGRSYKGVSTIKSGIKQCPWSSMIPHEH-----SYRGKDLQENYCRNP 175
QY 181 RGEGBGPMCTSNPEVREYECDDIPQSEVECMTCNGESYRGKDLMDHTSSGKICQRMWDQTP 240
   |||||||
DB 176 RGEGBGPMCTSNPEVREYECDDIPQSEVECMTCNGESYRGKDLMDHTSSGKICQRMWDQTP 235
QY 241 HRHKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDPTRMWEYCAIKTCADNTMDTDVPL 300
   |||||||
DB 236 HRHKFLPERYPDKGFDNRYCRNPDGQPRPWCYTLDPTRMWEYCAIKTCADNTMDTDVPL 295
QY 301 ETTECIOGQGBGYRGVNTWNGIIPCQRMDSQYRHEHDMTPENFKCKDLRENYCRNPDGS 360
   |||||||
DB 296 ETTECIOGQGBGYRGVNTWNGIIPCQRMDSQYRHEHDMTPENFKCKDLRENYCRNPDGS 355
QY 361 ESPMCFETTDNIRVGYCSQIPNCDMSHGQDCYRGNGKNYGNLSQTSRGLTCSMWDKME 420
   |||||||
DB 356 ESPMCFETTDNIRVGYCSQIPNCDMSHGQDCYRGNGKNYGNLSQTSRGLTCSMWDKME 415
QY 421 DLHRHIFWEPDASKLNEYCRNPDDDAHGPMCYTGNPLIPDYCPISRCBEDTPTIYNL 480
   |||||||
DB 416 DLHRHIFWEPDASKLNEYCRNPDDDAHGPMCYTGNPLIPDYCPISRCBEDTPTIYNL 475

```

QY 481 DHPVISCATKOLRVNNGIPTRTNIGMVSRLRNNKHCIGSGLIKESWVLTARQCFPSRD 540
 DB 476 DHPVISCATKOLRVNNGIPTRTNIGMVSRLRNNKHCIGSGLIKESWVLTARQCFPSRD 535
 QY 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDVLMLKARAVLDDFVSTIDLP 600
 DB 536 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDVLMLKARAVLDDFVSTIDLP 595
 QY 601 NYGCTIPEKTSYVGGYGTGLINYDGLRLVAHLIYINGNEKCSQHRGKVTLNESEICAG 660
 DB 596 NYGCTIPEKTSYVGGYGTGLINYDGLRLVAHLIYINGNEKCSQHRGKVTLNESEICAG 655
 QY 661 AEKIGSPGCGDYGGLVCSQHKRMVLYIVPGRGCAIPRRPGEFVAVAYAKWHKII 720
 DB 656 AEKIGSPGCGDYGGLVCSQHKRMVLYIVPGRGCAIPRRPGEFVAVAYAKWHKII 715
 QY 721 LTYKVPOS 728
 DB 716 LTYKVPOS 723

RESULT 10
 US-09-194-326-2
 : Sequence 2, Application US/09194326
 : Patent No. 6306827

: GENERAL INFORMATION:
 : APPLICANT: Kinoshita, Masahiko
 : APPLICANT: Ogawa, Hiroki
 : APPLICANT: Masanaga, Hiroaki
 : APPLICANT: Kobayashi, Fumie
 : APPLICANT: Yamaguchi, Kyoji
 : APPLICANT: Higashio, Kanji
 : TITLE OF INVENTION: Agent for Preventing and/or Treating Renal Disease
 : FILE REFERENCE: FJN-069
 : CURRENT APPLICATION NUMBER: US/09/194,326
 : CURRENT FILING DATE: 1998-11-24
 : EARLIER APPLICATION NUMBER: PCT/JP98/01221
 : EARLIER FILING DATE: 1998-03-20
 : EARLIER APPLICATION NUMBER: JP 94989
 : EARLIER FILING DATE: 1997-03-28
 : NUMBER OF SEQ ID NOS: 3
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 2
 : LENGTH: 723
 : TYPE: PRT
 : ORGANISM: Artificial Sequence
 : FEATURE:
 : OTHER INFORMATION: Description of Artificial Sequence: RRRR2AAA
 : OTHER INFORMATION: mutant of TCF-II
 : 09-194-326-2

Query Match 98.5%; Score 4064.5; DB 4; Length 723;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 719; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

QY 1 MMVTKLLPALLQVHLHLLLPLAIPYAEGRKRRNTIHEFFKSATKTLKIDPALKIK 60
 DB 1 MMVTKLLPALLQVHLHLLLPLAIPYAEGRKRRNTIHEFFKSATKTLKIDPALKIK 60
 QY 61 TKKVTADQCANRCTRNKGLPFTCKAFVDRKARKOCLMFPNSMSSGVKKEGHEFDLYE 120
 DB 61 TKKVTADQCANRCTRNKGLPFTCKAFVDRKARKOCLMFPNSMSSGVKKEGHEFDLYE 120
 QY 121 NKDYIRNCITIGKRSYKTVSITKSGIKCQPMSSMIPEHSHFLPSYSGKDLQENYCNP 180
 DB 121 NKDYIRNCITIGKRSYKTVSITKSGIKCQPMSSMIPEHSHFLPSYSGKDLQENYCNP 175
 QY 181 RGEESGPGCFTSNPEVREYVDIPCCSEVECTNGESYRGIMDHTESGKICQARDHOTP 240
 DB 176 RGEESGPGCFTSNPEVREYVDIPCCSEVECTNGESYRGIMDHTESGKICQARDHOTP 235
 QY 241 HRHKLPERYPDKGFDDNYCRNPDQPRPWCYTLDPHTRWEXCAIKTCADNTMNDTIVPL 300

DB 236 HRHKLPERYPDKGFDDNYCRNPDQPRPWCYTLDPHTRWEXCAIKTCADNTMNDTIVPL 295
 QY 301 ETTCTIGGEGYRGYVNTIWNGLPCQRMDSQYRPHEDMTPENFCKCKDLRENTCRNPDGS 360
 DB 296 ETTCTIGGEGYRGYVNTIWNGLPCQRMDSQYRPHEDMTPENFCKCKDLRENTCRNPDGS 355
 QY 361 ESPMCFETDINIRVYCSQIPNCDMSHGDCYRGNKNGYNTSQRSGGLTCSMMDK NME 420
 DB 356 ESPMCFETDINIRVYCSQIPNCDMSHGDCYRGNKNGYNTSQRSGGLTCSMMDK NME 415
 QY 421 DLHRHIFWEPDASKLNNYCRNPDGDAHGPWCYTGNDLPMDCPISRCBGDTPTIVNL 480
 DB 416 DLHRHIFWEPDASKLNNYCRNPDGDAHGPWCYTGNDLPMDCPISRCBGDTPTIVNL 475
 QY 481 DHPVISCATKOLRVNNGIPTRTNIGMVSRLRNNKHCIGSGLIKESWVLTARQCFPSRD 540
 DB 476 DHPVISCATKOLRVNNGIPTRTNIGMVSRLRNNKHCIGSGLIKESWVLTARQCFPSRD 535
 QY 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDVLMLKARAVLDDFVSTIDLP 600
 DB 536 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDVLMLKARAVLDDFVSTIDLP 595
 QY 601 NYGCTIPEKTSYVGGYGTGLINYDGLRLVAHLIYINGNEKCSQHRGKVTLNESEICAG 660
 DB 596 NYGCTIPEKTSYVGGYGTGLINYDGLRLVAHLIYINGNEKCSQHRGKVTLNESEICAG 655
 QY 661 AEKIGSPGCGDYGGLVCSQHKRMVLYIVPGRGCAIPRRPGEFVAVAYAKWHKII 720
 DB 656 AEKIGSPGCGDYGGLVCSQHKRMVLYIVPGRGCAIPRRPGEFVAVAYAKWHKII 715
 QY 721 LTYKVPOS 728
 DB 716 LTYKVPOS 723

RESULT 11
 US-09-194-326-3
 : Sequence 3, Application US/09194326
 : Patent No. 6306827

: GENERAL INFORMATION:
 : APPLICANT: Kinoshita, Masahiko
 : APPLICANT: Ogawa, Hiroki
 : APPLICANT: Masanaga, Hiroaki
 : APPLICANT: Kobayashi, Fumie
 : APPLICANT: Yamaguchi, Kyoji
 : APPLICANT: Higashio, Kanji
 : TITLE OF INVENTION: Agent for Preventing and/or Treating Renal Disease
 : FILE REFERENCE: FJN-069
 : CURRENT APPLICATION NUMBER: US/09/194,326
 : CURRENT FILING DATE: 1998-11-24
 : EARLIER APPLICATION NUMBER: PCT/JP98/01221
 : EARLIER FILING DATE: 1998-03-20
 : EARLIER APPLICATION NUMBER: JP 94989
 : EARLIER FILING DATE: 1997-03-28
 : NUMBER OF SEQ ID NOS: 3
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 3
 : LENGTH: 723
 : TYPE: PRT
 : ORGANISM: Artificial Sequence
 : FEATURE:
 : OTHER INFORMATION: Description of Artificial Sequence: KIKTKK27AIAVAA
 : OTHER INFORMATION: mutant of TCF-II
 : US-09-194-326-3

Query Match 98.5%; Score 4064.5; DB 4; Length 723;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 719; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

QY 1 MMVTKLLPALLQVHLHLLLPLAIPYAEGRKRRNTIHEFFKSATKTLKIDPALKIK 60
 DB 1 MMVTKLLPALLQVHLHLLLPLAIPYAEGRKRRNTIHEFFKSATKTLKIDPALKIK 60

Db 1 MWTKLLPALLQHVLLHLLPLLAIPYAEQORRRNTIHEFKSAKTLIKIDPALAIA 60
QY 61 TKKVTADOCANRCRTRNGKLPFTCKAFVFDKARKOCLMFPPNSMSGYKKEGHEFDLYE 120
Db 61 TAAVNTADOCANRCRTRNGKLPFTCKAFVFDKARKOCLMFPPNSMSGYKKEGHEFDLYE 120
QY 121 NKDYIRNCIIGKRSYKGVSTVITSGIKQPMSSMIPEHNSFLPSSYRGKDLQENYCRNP 180
Db 121 NKDYIRNCIIGKRSYKGVSTVITSGIKQPMSSMIPEHNSFLPSSYRGKDLQENYCRNP 175
QY 181 RGEESGPMCFSTSNPEVREVDIPQCSFEVCEMTGCESYRGMLMHTESGKICQRMDHOTP 240
Db 176 RGEESGPMCFSTSNPEVREVDIPQCSFEVCEMTGCESYRGMLMHTESGKICQRMDHOTP 235
QY 241 HRHFLPERYDCKGFDNDYCRNPDPGQPRPWCYTLDPHTRWEYCAIKTKADTMDNDVPL 300
Db 236 HRHFLPERYDCKGFDNDYCRNPDPGQPRPWCYTLDPHTRWEYCAIKTKADTMDNDVPL 295
QY 301 ETTECIGQSGEGYRGTVNTIWNIGIPQQRWDSQYRPHEDMTPEHNFCKDLRENYCRNPDS 360
Db 296 ETTECIGQSGEGYRGTVNTIWNIGIPQQRWDSQYRPHEDMTPEHNFCKDLRENYCRNPDS 355
QY 361 ESPMCFETDPMIRVGYCSQIIPNCMSHGQDCYRGNGKNYMNLSQTRSGLTCSMDKXME 420
Db 356 ESPMCFETDPMIRVGYCSQIIPNCMSHGQDCYRGNGKNYMNLSQTRSGLTCSMDKXME 415
QY 421 DLHRIEWEPPASKLNEYCRNPDDAHPWCYTGPNPLIPMDYCIISRCEGDTPTIYNL 480
Db 416 DLHRIEWEPPASKLNEYCRNPDDAHPWCYTGPNPLIPMDYCIISRCEGDTPTIYNL 475
QY 481 DHPVISCATKQRLRVNNGIPTRTNIGMWSLRYNKHAICGSLIKESWVLTARQCFPSRD 540
Db 476 DHPVISCATKQRLRVNNGIPTRTNIGMWSLRYNKHAICGSLIKESWVLTARQCFPSRD 535
QY 541 LKQDEAWLGIHDVHGRGDEKCKQVNLVNSQVLYGREGSDVLYMLKLARAVLDVFTIDLP 600
Db 536 LKQDEAWLGIHDVHGRGDEKCKQVNLVNSQVLYGREGSDVLYMLKLARAVLDVFTIDLP 595
QY 601 NYGCTIPKETSQSYGMYGTGLINYGILRVANLXIMGENKCSQHHRRKVTLLNESEICAG 660
Db 596 NYGCTIPKETSQSYGMYGTGLINYGILRVANLXIMGENKCSQHHRRKVTLLNESEICAG 655
QY 661 AEKIGSGCEBDYGGPLVCEQHKMRYLVGIVPGRGCAIPNRPGLFVNAVAYAKMIHKII 720
Db 656 AEKIGSGCEBDYGGPLVCEQHKMRYLVGIVPGRGCAIPNRPGLFVNAVAYAKMIHKII 715
QY 721 LTYKVPQS 728
Db 716 LTYKVPQS 723

RESULT 12
: Sequence 50, Application US/08460890A
: Patent No. 5994109
: GENERAL INFORMATION:
: APPLICANT: Moo, Savio L.C.
: APPLICANT: Smith, Louis C.
: APPLICANT: Cristiano, Richard J.
: APPLICANT: Gotcheak, Stephen
: TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
: NUMBER OF SEQUENCES: 65
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: STREET: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,890A
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 697 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-890A-50

Query Match 96.2%; Score 3968; DB 2; Length 697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 ORKRRNTIHEFKSAKTLIKIDPALKIKTKKVNADOCANRCRTRNGKLPFTCKAFVFDK 91
Db 1 ORKRRNTIHEFKSAKTLIKIDPALKIKTKKVNADOCANRCRTRNGKLPFTCKAFVFDK 60
QY 92 ARKQCLMFPNMSGSGYKKEGHEFDLYENKDYIRNCIIGKRSYKGVSTVITSGIKQCP 151
Db 61 ARKQCLMFPNMSGSGYKKEGHEFDLYENKDYIRNCIIGKRSYKGVSTVITSGIKQCP 120
QY 152 WSSMTPEHSHFLPSSYRKRDQENYCRNPREGGPMCFSTSNPEVREVCDDIPQSEVEC 211
Db 121 WSSMTPEHSHFLPSSYRKRDQENYCRNPREGGPMCFSTSNPEVREVCDDIPQSEVEC 180
QY 212 MTCNGESYRGMLDHTESGKICQRMHDQTPHRRKFLPERYPDKGPDNCRNPDPGQPRPWC 271
Db 181 MTCNGESYRGMLDHTESGKICQRMHDQTPHRRKFLPERYPDKGPDNCRNPDPGQPRPWC 240
QY 272 YTLDPHTRWEYCAIKTKADNTMDNDVPLETTECIGQSGEGYRGTVNTIWNIGIPQQRWDS 331
Db 241 YTLDPHTRWEYCAIKTKADNTMDNDVPLETTECIGQSGEGYRGTVNTIWNIGIPQQRWDS 300
QY 332 QYRPHEDMTPEHNFCKDLRENYCRNPDSSEPMCTTPDNIRVGYCSQIPNCDSHGQDC 391
Db 301 QYRPHEDMTPEHNFCKDLRENYCRNPDSSEPMCTTPDNIRVGYCSQIPNCDSHGQDC 360
QY 392 YGNGKNYMNLSQTRSGLTCSMDKXMEDLHRIHFWERDASKLNEYCRNPDDAHPW 451
Db 361 YGNGKNYMNLSQTRSGLTCSMDKXMEDLHRIHFWERDASKLNEYCRNPDDAHPW 420
QY 452 CYTGNPLIPMDYCIISRCEGDTPTIYNLDHPVISCATKQRLRVNNGIPTRTNIGMWSL 511
Db 421 CYTGNPLIPMDYCIISRCEGDTPTIYNLDHPVISCATKQRLRVNNGIPTRTNIGMWSL 480
QY 512 RYRNKHIGGSLIKESWVLTARQCFPSRDLDYEWMLGIHDVHGRGDEKCKQVNLVNSQV 571
Db 481 RYRNKHIGGSLIKESWVLTARQCFPSRDLDYEWMLGIHDVHGRGDEKCKQVNLVNSQV 540

QY 572 YGPESSDLVLMKLARPAVLDFVSTIDLPNGCTIPEKTSVSGVGMGTGLINDGLLRV 631
DB 541 YGPESSDLVLMKLARPAVLDFVSTIDLPNGCTIPEKTSVSGVGMGTGLINDGLLRV 600
QY 632 AHLTYMGNEKCSQHHRGKVTTLNESEICAGAEKIGSGPEGDYGGLVCEQHMRVNLGYI 691
DB 601 AHLTYMGNEKCSQHHRGKVTTLNESEICAGAEKIGSGPEGDYGGLVCEQHMRVNLGYI 660
QY 692 YGRCALPNNRPGIFVRAVAYAKWIKIILTYKVPQS 728
DB 661 YGRCALPNNRPGIFVRAVAYAKWIKIILTYKVPQS 697

RESULT 13
US-08-167-641C-50
; Sequence 50, Application US/08167641C
; Patent No. 6033884
; GENERAL INFORMATION:
; APPLICANT: WOO, Saviio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristliano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08167,641C
; FILING DATE: December 14, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/855,389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 205/012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 697 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-167-641C-50

Query Match 96.2%; Score 3968; DB 3; Length 697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 QRRRRNTIHEFKSAKTTLIKIDPALKIKTKKVVNTADOCANRCKTRNGKGLPFTCKAVFEDK 91
DB 1 QRRRRNTIHEFKSAKTTLIKIDPALKIKTKKVVNTADOCANRCKTRNGKGLPFTCKAVFEDK 60

QY 92 ARKOCLEPFENSMSSGVAKKEGHEFDLYENKDYIRNCLIGKSGKSTVSTTKGICOP 151
DB 61 ARKOCLEPFENSMSSGVAKKEGHEFDLYENKDYIRNCLIGKSGKSTVSTTKGICOP 120
QY 152 WSMIPEHSHFLPSSYRKDLQENYCRNPGEEGPWCFTSNPEVREVCIDIPGCSVEVC 211
DB 121 WSMIPEHSHFLPSSYRKDLQENYCRNPGEEGPWCFTSNPEVREVCIDIPGCSVEVC 180
QY 212 MTCNGESYRGLMDHTESGKIQARDHQTPIHRKFLPERYPDKGFDNYCRNPDGQPRWC 271
DB 181 MTCNGESYRGLMDHTESGKIQARDHQTPIHRKFLPERYPDKGFDNYCRNPDGQPRWC 240
QY 272 YTLDPHTWMEYCAIKTCADNTMNDTVPLETTTECIGOGGEGYRGVTNIMNGIPCOMRDS 331
DB 241 YTLDPHTWMEYCAIKTCADNTMNDTVPLETTTECIGOGGEGYRGVTNIMNGIPCOMRDS 300
QY 332 QYPHEHDMTPENFKCKDLRENYCRNPDSGSPWCFTTDPNIRVGCQIIPNCDMSHGDC 391
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDSGSPWCFTTDPNIRVGCQIIPNCDMSHGDC 360
QY 392 YRGNKNYMGNLSTQSGLTCSMDKXNEDLHRHIFWEPDASKLINENYCRNPDDDAHGPW 451
DB 361 YRGNKNYMGNLSTQSGLTCSMDKXNEDLHRHIFWEPDASKLINENYCRNPDDDAHGPW 420
QY 452 CYTGNPILPMDYCPISRCGDTPTIVNLDHPVISCATKOLRVNGIPTRTNIGMNVSL 511
DB 421 CYTGNPILPMDYCPISRCGDTPTIVNLDHPVISCATKOLRVNGIPTRTNIGMNVSL 480
QY 512 RYRNKHICGSLIKESWVLTAOCPPSRDLDEYAMLGIHDVHGGRDECKOVLNVSLV 571
DB 481 RYRNKHICGSLIKESWVLTAOCPPSRDLDEYAMLGIHDVHGGRDECKOVLNVSLV 540
QY 572 YGPESSDLVLMKLARPAVLDFVSTIDLPNGCTIPEKTSVSGVGMGTGLINDGLLRV 631
DB 541 YGPESSDLVLMKLARPAVLDFVSTIDLPNGCTIPEKTSVSGVGMGTGLINDGLLRV 600
QY 632 AHLTYMGNEKCSQHHRGKVTTLNESEICAGAEKIGSGPEGDYGGLVCEQHMRVNLGYI 691
DB 601 AHLTYMGNEKCSQHHRGKVTTLNESEICAGAEKIGSGPEGDYGGLVCEQHMRVNLGYI 660
QY 692 YGRCALPNNRPGIFVRAVAYAKWIKIILTYKVPQS 728
DB 661 YGRCALPNNRPGIFVRAVAYAKWIKIILTYKVPQS 697

RESULT 14
US-08-460-971A-50
; Sequence 50, Application US/08460971A
; Patent No. 6150168
; GENERAL INFORMATION:
; APPLICANT: WOO, Saviio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristliano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/460,971A
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/063
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 697 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-971A-50

Query Match 96.2%; Score 3968; DB 4; Length 697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 ORRRNTIHEFKSATTLLIKIDPALIKITKKVNTADOCANRCKTRNGGLPTCKAFVFDK 91
DB 1 ORRRNTIHEFKSATTLLIKIDPALIKITKKVNTADOCANRCKTRNGGLPTCKAFVFDK 60
QY 92 ARKOCLEFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKTVSITKSGIKOP 151
DB 61 ARKOCLEFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKTVSITKSGIKOP 120
QY 152 WSMITHFHSFLPSSYRGKDLQENYCRNPRGEGGPMWCTSNPEVREYVCDIPQCSVEEC 211
DB 121 WSMITHFHSFLPSSYRGKDLQENYCRNPRGEGGPMWCTSNPEVREYVCDIPQCSVEEC 180
QY 212 MTCNGSESYGLMHTSGKICQWMDHQTPIRHRKFLPERYPDKGFDNDYCNPPGQPPRWC 271
DB 181 MTCNGSESYGLMHTSGKICQWMDHQTPIRHRKFLPERYPDKGFDNDYCNPPGQPPRWC 240
QY 272 YLLDPTRMFYCAIKTCADNTMTDVTPLTECTICGQGGEGYGVNTINNGIIPCQWRMS 331
DB 241 YLLDPTRMFYCAIKTCADNTMTDVTPLTECTICGQGGEGYGVNTINNGIIPCQWRMS 300
QY 332 QYRHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCDSHGQDC 391
DB 301 QYRHEHDMTPENFKCDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCDSHGQDC 360
QY 392 YRNGKNYKGNLSQTSRGLTCSMWDKMMEDLHRHIEWEPASKLNEYCNPPDDAHPW 451
DB 361 YRNGKNYKGNLSQTSRGLTCSMWDKMMEDLHRHIEWEPASKLNEYCNPPDDAHPW 420
QY 452 CYTGNPLIPMDYCPISRCEDGTPPTIVNDHPYISCAKTKQLRVNGLIPRTNIGMWVSL 511
DB 421 CYTGNPLIPMDYCPISRCEDGTPPTIVNDHPYISCAKTKQLRVNGLIPRTNIGMWVSL 480
QY 512 RYRNKHCGLSLIKESWVLTAROCFSPRDLKDYEAWLGIDHVGRODEKCKOVLNVSOVL 571
DB 481 RYRNKHCGLSLIKESWVLTAROCFSPRDLKDYEAWLGIDHVGRODEKCKOVLNVSOVL 540
QY 572 YRGGESDLYMLKARPAVLDVSTIDLRYGCTIPKTSQSVYGMKYTGLINDELRLV 631
DB 541 YRGGESDLYMLKARPAVLDVSTIDLRYGCTIPKTSQSVYGMKYTGLINDELRLV 600
QY 632 ARLYINGNEKCSQHRKQVTLNSESICAGAEKIGSGPCEDGYGGLVCEQHKRMVGLVI 691

DB 601 ARLYINGNEKCSQHRKQVTLNSESICAGAEKIGSGPCEDGYGGLVCEQHKRMVGLVI 660
QY 692 VPRGCAIPNRPGLFVAVAYYAKVHKIILTYKVPQS 728
DB 661 VPRGCAIPNRPGLFVAVAYYAKVHKIILTYKVPQS 697

RESULT 15
US-08-462-040-50
Sequence 50, Application US/08462040
Patent No. 6177554
GENERAL INFORMATION:
APPLICANT: WOO, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,040
FILING DATE: June 5, 1995
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992

APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993

ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/078

TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 697 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-462-040-50

Query Match 96.2%; Score 3968; DB 4; Length 697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 ORRRNTIHEFKSATTLLIKIDPALIKITKKVNTADOCANRCKTRNGGLPTCKAFVFDK 91
DB 1 ORRRNTIHEFKSATTLLIKIDPALIKITKKVNTADOCANRCKTRNGGLPTCKAFVFDK 60
QY 92 ARKOCLEFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKTVSITKSGIKOP 151

Db 61 ARKQCLMFFPNSMSGVKKEFGHEFDLYENKDYIRNCLIGKGRYKGVSTIKSGIKOP 120
QY 152 WSSMIPHEHSPLSSSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDIPOCSEVEC 211
Db 121 WSSMIPHEHSPLSSSYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDIPOCSEVEC 180
QY 212 MTCNGESYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDIPOCSEVEC 271
Db 181 MTCNGESYRGKDLQENYCRNPRGEGGPMCTSNPEVRYEVCDIPOCSEVEC 240
QY 272 YTLDPHTRMWYCAIKTCADFTMNDTVPLETTTCIOGOGEGYRGVNTIWNIGIPCORWDS 331
Db 241 YTLDPHTRMWYCAIKTCADFTMNDTVPLETTTCIOGOGEGYRGVNTIWNIGIPCORWDS 300
QY 332 QYHEHDMTPENFKCKDLRENYCRNPDGSESPMCTTDPNIRVGYCSQIIPNCDMSHGQDC 391
Db 301 QYHEHDMTPENFKCKDLRENYCRNPDGSESPMCTTDPNIRVGYCSQIIPNCDMSHGQDC 360
QY 392 YRGNGKNYMGNLQOTRSGLTCSMDKNMEDLHRHIFWEPDASKLNEYCRNPDGDAHGPW 451
Db 361 YRGNGKNYMGNLQOTRSGLTCSMDKNMEDLHRHIFWEPDASKLNEYCRNPDGDAHGPW 420
QY 452 CYTGNPLIPMDYCPISRCEDGDTPTIYNLDHPYISCAKTKQLRVNGIPTRTNIGMWVSL 511
Db 421 CYTGNPLIPMDYCPISRCEDGDTPTIYNLDHPYISCAKTKQLRVNGIPTRTNIGMWVSL 480
QY 512 RYRNKHICGSSLIKESVNLTAROCPSRDLKDYEAWLGIHDVHGRGDECKOYLVNSQV 571
Db 481 RYRNKHICGSSLIKESVNLTAROCPSRDLKDYEAWLGIHDVHGRGDECKOYLVNSQV 540
QY 572 YGPGSDLVLMKLARPAVLDDFVSTIDLPNYGCTIPEKTSQSVYGMGYTGLINYDGLRV 631
Db 541 YGPGSDLVLMKLARPAVLDDFVSTIDLPNYGCTIPEKTSQSVYGMGYTGLINYDGLRV 600
QY 632 AHLIYMGNEKCSQHHRGKVTLNESEICAGAEKIGSPCEGDYGGPLYCEQHKMRMVLGYI 691
Db 601 AHLIYMGNEKCSQHHRGKVTLNESEICAGAEKIGSPCEGDYGGPLYCEQHKMRMVLGYI 660
QY 692 VPGRCALPNNRPGIFVRVAYYAKWIKHILLYKVPQS 728
Db 661 VPGRCALPNNRPGIFVRVAYYAKWIKHILLYKVPQS 697

Search completed: June 18, 2002, 17:47:21
Job time: 242 sec